GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: May 7, 2003, 15:14:59 ; Search time 43 seconds
(without alignments)
6273.536 Million cell updates/sec

Title: US-09-965-830-1_COPY_6_3257
Perfect score: 6089
Sequence: 1 atgcgcggcaatgcgggcct.........aagaaggcacaggggtctga 3252

Searched: 112892 seqs, 41476328 residues

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Total number of hits satisfying chosen parameters: 225784

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+_n2p.model -DEV=xlp
-Q=/cgn2_1/USPTO_spool/US0996583O/runat_07052003_151448_6247/app_query.fasta_1.3399
-DB=SwissProt_40 -QFMT=fastan -SUFFIX=rsp -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=-1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US0996583O.@CGN_1.1_70.@runat_07052003_151448_6247 -NCPU=6 -ICPU=3
-NO_XLPXV -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOG -DEV_TIMEOUT=120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1551	25.5	1159	1	HERG_HUMAN	homo sapien
2	1390	22.8	1174	1	CIKE_DROME	drosophila
3	439	7.2	706	1	CNG3_BOVIN	bos taurus
4	429.5	7.1	683	1	CNG1_RAT	r cgmp-gate
5	424.5	7.0	694	1	CNG3_HUMAN	homo sapien
6	423.5	7.0	735	1	CNG1_CHICK	gallus gall
7	422.5	6.9	631	1	CNG3_MOUSE	mus musculu
8	421.5	6.9	665	1	CNG_DROME	drosophila
9	421.5	6.9	691	1	CNG1_CANFA	c cgmp-gate
10	419	6.9	684	1	CNG1_MOUSE	m cgmp-gate
11	414.5	6.8	682	1	CNG1_ICTPU	ictalurus p
12	413.5	6.8	690	1	CNG1_BOVIN	bos taurus
13	412	6.8	686	1	CNG1_HUMAN	homo sapien
14	411	6.7	645	1	CNG3_CHICK	gallus gall
15	405	6.7	664	1	CNG2_RABIT	oryctolagus
16	397.5	6.7	1464	1	CA11_HUMAN	homo sapien
17	395	6.5	664	1	CNG2_MOUSE	mus musculu
18	393	6.5	664	1	CNG2_RAT	rattus norv

ALIGNMENTS

Result No.	Score	% Query Match	Length	DB	ID	AC	Description
19	391.5	6.4	663	1	CNG2_BOVIN	Q03041	bos taurus
20	391	6.4	1453	1	CA11_MOUSE	P11087	mus musculu
21	382	6.3	1460	1	CA11_CANFA	Q9xsj7	canis famil
22	379	6.3	1453	1	CA11_CANFA	Q9xsj7	canis famil
23	378	6.3	1460	1	CA11_CANFA	Q9xsj7	canis famil
24	369.5	6.2	1453	1	CNGX_RAT	P24457	gallus gall
25	365	6.1	575	1	CA11_CHICK	Q64359	rattus norv
26	360.5	6.0	1464	1	CA11_CHICK	P24452	homo sapien
27	358.5	6.0	1453	1	CA13_CHICK	P12105	gallus gall
28	358	5.9	1355	1	CA11_MOUSE	P24457	gallus gall
29	358	5.9	1262	1	CA21_RANCA	P11087	mus musculu
30	352.5	5.8	1364	1	CA21_BOVIN	O42350	rana catesb
31	350	5.8	1356	1	CA21_ONCMY	P02465	bos taurus
32	348	5.7	1049	1	CA13_BOVIN	O93484	oncorhynchu
33	346.5	5.7	1806	1	CA1B_HUMAN	P04258	bos taurus
34	346	5.7	1805	1	CA1B_HUMAN	P12107	homo sapien
35	346	5.7	1418	1	CA1B_MOUSE	P02458	homo sapien
36	346	5.7	1804	1	CNG_CAEEL	O61245	mus musculu
37	344	5.6	733	1	CNG_CAEEL	O03611	caenorhabdi
38	342.5	5.6	1464	1	CA13_MOUSE	P08121	mus musculu
39	341.5	5.6	1466	1	CA13_HUMAN	P08121	mus musculu
40	340.5	5.6	779	1	CA11_BOVIN	P02461	homo sapien
41	338.5	5.6	1459	1	CA12_MOUSE	P02453	bos taurus
42	336.5	5.5	1262	1	CA13_CHICK	P28481	mus musculu
43	333.5	5.5	1366	1	CA21_CANFA	P12105	gallus gall
44	332	5.5	1255	1	MUC1_HUMAN	O46392	canis famil
45	331.5	5.5	1366	1	CA21_HUMAN	P15941	h mucin 1 p
	330	5.4	1372	1	CA21_RAT	P08123	homo sapien
						P02466	rattus norv

RESULT 1
HERG_HUMAN

ID   HERG_HUMAN   STANDARD;   PRT;   1159 AA.
AC   Q12809; O75680;
DT   30-MAY-2000 (Rel. 39, Created)
DT   30-MAY-2000 (Rel. 39, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Voltage-gated potassium channel HERG (Potassium channel, voltage-gated
DE   subfamily H member 2) (Ether-a-go-go-related protein).
GN   KCNH2 OR HERG.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Hippocampus;
RX   MEDLINE=94211879; PubMed=8159766;
RA   Warmke J.W., Ganetzky B.;
RT   "A family of potassium channel genes related to eag in Drosophila and
RT   mammals.";
RL   Proc. Natl. Acad. Sci. U.S.A. 91:3438-3442(1994).
RN   [2]
RP   SEQUENCE FROM N.A., AND VARIANT LQT2 CYS-534.
RX   MEDLINE=98260867; PubMed=9600240;
RA   Itoh T., Tanaka T., Nagai R., Kamiya T., Sawayama T., Nakayama T.,
RA   Tomoike H., Sakurada H., Yazaki Y., Nakamura Y.;
RT   "Genomic organization and mutational analysis of HERG, a gene
RT   responsible for familial long QT syndrome.";
RL   Hum. Genet. 102:435-439(1998).
RN   [3]
RP   SEQUENCE OF 27-1159 FROM N.A.
RA   Yang P., Paulussen A., Verhasselt P., Crabbe R., Luyten W.,
RA   Armstrong M.;
RT   "Analysis of the human HERG gene: intron localisation and
RT   identification of a novel inherited mutation associated with long
RT   QT.";
RL   Submitted (AUG-1998) to the EMBL/GenBank/DDBJ databases.
RN   [4]
RP   VARIANTS LQT2 ASP-470; VAL-561; SER-628 AND 500-ILE--PHE-508 DEL.
RX   MEDLINE=95196272; PubMed=7889573;
RA   Curran M.E., Splawski I., Timothy K.W., Vincent G.M., Green E.D.,

RA Keating M.T.;
RT "A molecular basis for cardiac arrhythmia: HERG mutations cause long
RT QT syndrome.";
RL Cell 80:795-803(1995).
RN [5]
RP VARIANT LQT2 MET-822.
RX MEDLINE=97071892; PubMed=8914737;
RA Benson D.W., MacRae C.A., Vesely M.R., Walsh E.P., Plummer M.H., Ginsburg G.S.,
RA Satler C.A., Walsh E.P., Vesely M.R., Plummer M.H., Ginsburg G.S.,
RA Seidman C.E., Satler C.A.,
RA Jacob H.J.;
RT "Novel missense mutation in the cyclic nucleotide-binding domain of
RT HERG causes long QT syndrome.";
RL Am. J. Med. Genet. 65:27-35(1996).
RN [6]
RP VARIANT LQT2 ARG-593.
RX MEDLINE=96259954; PubMed=8635257;
RA Tanaka T., Nagai R., Tomoike H., Takata S., Yano K., Yabuta K.,
RA Haneda N., Nakano O., Shibata A., Sawayama T., Kasai H., Yazaki Y.,
RA Nakamura Y.;
RT "Four novel KVLQT1 and four novel HERG mutations in familial long-QT
RT syndrome.";
RL Circulation 95:565-567(1997).
RN [7]
RP VARIANTS LQT2 ILE-474; VAL-561; HIS-611; VAL-614 AND ALA-630.
RX MEDLINE=97176600; PubMed=9024139;
RA Splawski I., Shen J., Timothy K.W., Vincent G.M., Lehmann M.H.,
RA Keating M.T.;
RT "Genomic structure of three long QT syndrome genes: KVLQT1, HERG, and
RT KCNE1.";
RL Genomics 51:86-97(1998).
RN [8]
RP VARIANTS LQT2 CYS-572; ASP-588; VAL-614; ASP-629; SER-629 AND SER-633.
RX MEDLINE=98360095; PubMed=9693036;
RA Satler C.A., Vesely M.R., Duggal P., Ginsburg G.S., Beggs A.H.;
RT "Multiple different missense mutations in the pore region of HERG in
RT patients with long QT syndrome.";
RL Hum. Genet. 102:265-272(1998).
RN [9]
RP VARIANTS LQT2 LEU-612; VAL-614; ASP-629; SER-629 AND SER-633.
RX MEDLINE=98204397; PubMed=9544837;

RN [10]
RP VARIANTS LQT2 PRO-558; CYS-582; SER-604; MET-613 AND LEU-640.
RX MEDLINE=99235550; PubMed=10220144;
RA Jongbloed R.J.E., Wilde A.A.M., Geelen J.L.M.C., Doevendans P.,
RA Schaap C., Van Langen I., van Tintelen J.P., Cobben J.M.,
RA Beaufort-Krol G.C.M., Geraedts J.P.M., Smeets H.J.M.;
RT "Novel KCNQ1 and HERG missense mutations in Dutch long-QT families.";
RL Hum. Mutat. 13:301-310(1999).
RN [11]
RP VARIANTS LQT2 SER-601.
RX MEDLINE=99112459; PubMed=9452080;
RA Akimoto K., Furutani M., Imamura S.-I., Furutani Y., Kasanuki H.,
RA Takao A., Momma K., Matsuoka R.;
RT "Novel missense mutation (G601S) of HERG in a Japanese long QT
RT syndrome family.";
RL Hum. Mutat. Suppl. 1:S184-S186(1998).
RN [12]
RP VARIANT LQT2 ARG-572.
RX MEDLINE=99235552; PubMed=10220146;
RA Larsen L.A., Christiansen M., Vuust J., Andersen P.S.;
RT "High-throughput single-strand conformation polymorphism analysis by
RT automated capillary electrophoresis: robust multiplex analysis and
RT pattern-based identification of allelic variants.";
RL Hum. Mutat. 13:318-327(1999).
RN [13]
RP VARIANTS LQT2 L-29; T-33; R-53; Q-56; G-66; R-70; P-78 AND R-86.
RX MEDLINE=99214568; PubMed=10187793;
RA Chen J., Zou A., Splawski I., Keating M.T., Sanguinetti M.C.;
RT "Long QT syndrome-associated mutations in the Per-Arnt-Sim (PAS)

RT domain of HERG potassium channels accelerate channel deactivation.";
RL J. Biol. Chem. 274:10113-10118(1999).
RN [14]
RP VARIANT LQT2 LYS-629.
RX MEDLINE=99445248; PubMed=10517660;
RA Yoshida H., Horie M., Otani H., Takano M., Tsuji K., Kubota T.,
RA Fukunami M., Sasayama S.;
RT "Characterization of a novel missense mutation in the pore of HERG in
RT a patient with long QT syndrome.";
RL J. Cardiovasc. Electrophysiol. 10:1262-1270(1999).
RN [15]
RP VARIANT LQT2 ARG-572.
RX MEDLINE=20197680; PubMed=10735633;
RA Larsen L.A., Svendsen I.H., Jensen A.M., Kanters J.K., Andersen P.S.,
RA Moller M., Sorensen S.A., Sandoe E., Jacobsen J.R., Vuust J.,
RA Christiansen M.;
RT "Long QT syndrome with a high mortality rate caused by a novel G572R
RT missense mutation in KCNH2.";
RL Clin. Genet. 57:125-130(2000).
RN [16]
RP VARIANTS LQT2.
RX MEDLINE=20432616; PubMed=10973849;
RA Splawski I., Shen J., Timothy K.W., Lehmann M.H., Priori S.,
RA Robinson J.L., Moss A.J., Schwartz P.J., Towbin J.A., Vincent G.M.,
RA Keating M.T.;
RT "Spectrum of mutations in long-QT syndrome genes. KVLQT1, HERG, SCN5A,
RT KCNE1, and KCNE2.";
RL Circulation 102:1178-1185(2000).
-!- FUNCTION: INWARDLY RECTIFYING CARDIAC POTASSIUM (IKR) CHANNEL.
-!- SUBUNIT: MAY ASSEMBLE WITH EITHER MIRP1 OR MINK. COMPLEXES WITH
    MIRP1 ARE MORE STABLE.
-!- TISSUE SPECIFICITY: HIGHLY EXPRESSED IN HEART.
-!- DISEASE: DEFECTS IN KCNH2 ARE ASSOCIATED WITH LONG QT SYNDROME
    TYPE 2 (LQT2). THIS DISEASE IS CHARACTERIZED BY A PROLONGED QT
    INTERVAL ON THE ECG AND POLYMORPHIC VENTRICULAR ARRHYTHMIAS KNOWN
    AS TORSADES DE POINTES. THESE ARRHYTHMIAS OFTEN OCCUR IN RELATION
    TO EXERCISE OR EMOTIONAL STRESS AND MAY RESULT IN RECURRENT
    SYNCOPE, SEIZURES, OR SUDDEN CARDIAC DEATH. DEAFNESS IS OFTEN
    ASSOCIATED WITH THE SYNDROME.
-!- MISCELLANEOUS: IT HAS THE ARCHITECTURAL PLAN OF THE
    DEPOLARIZATION-ACTIVATED POTASSIUM CHANNEL FAMILY (6 PUTATIVE
    TRANSMEMBRANE SEGMENTS), YET IT EXHIBITS RECTIFICATION LIKE THAT
    OF THE INWARD-RECTIFYING POTASSIUM CHANNELS.
-!- SIMILARITY: CONTAINS 1 CYCLIC NUCLEOTIDE-BINDING DOMAIN.
-!- SIMILARITY: CONTAINS 1 PAS (PER-ARNT-SIM) DIMERIZATION DOMAIN.
-!- SIMILARITY: CONTAINS 1 PAS-ASSOCIATED C-TERMINAL (PAC) DOMAIN.
-!- DATABASE: NAME=LQTSdb; NOTE=KCNH2 mutations page;
    WWW="http://www.ssi.dk/en/forskning/lqtsdb/herg.htm".

DR EMBL; U04270; AAA62473.1; -.
DR EMBL; AB009071; BAA37096.1; -.
DR EMBL; AB009070; BAA37096.1; JOINED.
DR EMBL; AB009069; BAA37096.1; JOINED.
DR EMBL; AB009068; BAA37096.1; JOINED.
DR EMBL; AB009067; BAA37096.1; JOINED.
DR EMBL; AB009066; BAA37096.1; JOINED.
DR EMBL; AB009065; BAA37096.1; JOINED.
DR EMBL; AB009064; BAA37096.1; JOINED.
DR EMBL; AB009063; BAA37096.1; JOINED.
DR EMBL; AB009062; BAA37096.1; JOINED.
DR EMBL; AB009061; BAA37096.1; JOINED.
DR EMBL; AB009060; BAA37096.1; JOINED.
DR EMBL; AB009059; BAA37096.1; JOINED.
DR EMBL; AB009058; BAA37096.1; JOINED.
DR EMBL; AB009057; BAA37096.1; JOINED.

```
DR   EMBL; AJ010538; CAA09232.1; -
DR   EMBL; AJ010539; CAA09232.1; JOINED.
DR   EMBL; AJ010540; CAA09232.1; JOINED.
DR   EMBL; AJ010541; CAA09232.1; JOINED.
DR   EMBL; AJ010542; CAA09232.1; JOINED.
DR   EMBL; AJ010543; CAA09232.1; JOINED.
DR   EMBL; AJ010544; CAA09232.1; JOINED.
DR   EMBL; AJ010545; CAA09232.1; JOINED.
DR   EMBL; AJ010546; CAA09232.1; JOINED.
DR   EMBL; AJ010547; CAA09232.1; JOINED.
DR   EMBL; AJ010548; CAA09232.1; JOINED.
DR   EMBL; AJ010549; CAA09232.1; JOINED.
DR   EMBL; AJ010550; CAA09232.1; JOINED.
DR   EMBL; AJ010551; CAA09232.1; JOINED.
```

Alignment Scores:

```
Pred. No.:
Score:
Pred. No.:                 1.44e-53          Length:        1159
Score:                     1551.00           Matches:       424
Percent Similarity:        44.868            Conservative:  143
Best Local Similarity:     33.54%            Mismatches:    337
Query Match:               25.478            Indels:        360
                                             Gaps:          38
                           1                 DB:            1
```

US-09-965-830-1_COPY_6_3257 (1-3252) x HERG_HUMAN (1-1159)

```
Qy    1 ATGCCGCCGGCCANGCGGGCGGGGCCTCCTGCCGCCCTCAGACACCACCTTCCTGGACACCATCGGCTACG  60
Db    1        MetProValArgGlyHisLeuValAlaProGlnAsnThrPheLeuAspThrIleIleArg  20

Qy   61 CGCTTCGACGGCGACCACTAACTTCGTCGTGGGCANACGGCCCAGGTGGCGGGGCCCTTC  120
Db   21 LysPheGluGlyGlnSerArgLysPheIleIleAlaAsnAlaArgValGluAsnCys---  39

Qy  121 CCCGGTGGTCTACTGCTTCTGATGAGGGCTTCTGTTGACCTTCCACGGGGCTTCCCGGGCTCAGGTC  180
Db   40 AlaValIleTyrCysSerAsnAspGlyPheCysGlyLeuCysGlyTyrSerArgAlaGluVal  59

Qy  181 ATGCAGCGGGGCGTGCCCTGCTCCTTCCTTTATGGGCCAGACACCACCAGTGAGCTCGTGCGC  240
Db   60 MetGlnArgGlyCysThrCysSerAspPheLeuHisGlyCysProArgThrGlnArgArgAlaAla  79

Qy  241 CAACGAGATCGGCAAGGCCCTGGACGAGGCACAAGAGAGTTCAAAGGCTGAGCTGATCCTGTAC  300
Db   80 AlaGlnIleLeuAlaGlnAlaLeuLeuGlyAlaGluGlyAlaArgLysValGluIleAlaPheTyr  99

Qy  301 CGGAAGAGCGGGCGTTCCGTTCTGGTGTCTCCGGATGTGCGATACCCCATAAAGAATGAGAAA  360
Db  100 ArgLysAspGlySerCysSerPheLeuCysLeuValAspSerValValProGlyAlaValAsp  119

Qy  361 GGGGGAGGTGGGCTCTCCTA--  
Db  120 GlyAlaValIleMetPheIleLeuGlnAlaValMetGlyLysAspMetValGly  139

Qy  400 ------AGCGAAACCAAAGAACCGAGGGGGCCCCCGACGAGATGG----  435
Db  140 SerProAlaHisAlaSerpThrAsnHisArgGlyProProThrSerTrpLeuAlaProGlyArg  159

Qy  435 -------------  435
Db  160 AlaLysThrPheArgLeuLeuLysProAlaLeuLeuLeuThrAlaArgLeuSerSer  179

Qy  435 --AAGGAGACAGGTGGTGGCCGGCCGGCCGTATGGCCGGGC---ACGATCCAAAGGCTTC  489
Db  180 ValArgSerGlyAlaGlyAlaGly-GlyAlaGlyAla-ProGlyAlaGlyAlaValAspAlaVala  199

Qy  490 AATGCCAACCGGCCG----------  503
Db  199 spLeuThrProAlaAlaProSerGluSerLeuAlaLeuAspGluValThrAlaMeta  219

Qy  504 ---------CCGGACGCCGGCGGCCCGTGCTCTA-------  524
Db  219 spAsnHisValAlaGlyLeuGlyProAlaGluGluArgArgAlaLeuValGlyProGlyLys  239
```

Right-hand alignment column:

```
Qy  525 --CCACCTGTCCGGCACCTGCAGAAGCAGCAGCCCAAGGCAGAG---CACAAGCTCAAT---  576
Db  239 erProProArgSerAlaProGlyGlnLeu-ProSerProArgAlaHisSerLeuAsnPro  258

Qy  576 -------  576
Db  259 AspAlaSerGlySerSerCysSerLeuAlaArgThrArgGluSerCysAlaSer  278

Qy  577 -------  577
Db  279 ValArgArgAlaSerAlaAspIleGluAlaMetArgAlaGlyValLeuProPro  298

Qy  595 AAACCAAAC-------  603
Db  299 ProProArgHisAlaAlaSerThrGlyAlaMetHisProLeuArgSerGlyLeuAsnSer  318

Qy  603 -------  603
Db  319 ThrSerAspSerAspLeuValArgTyrArgThrIleSerLysIleProGlnIleThrLeu  338

Qy  603 -------  603
Db  339 AsnPheValAspLeuLysGlyAspProPheLeuAlaSerProThrSerAspArgGluIle  358

Qy  603 -------  603
Db  359 IleAlaProLysIleLysGluArgThrHisAsnValThrGluLysValThrGlnValLeu  378

Qy  604 ---------TTGCCTGAGTACAAAGTAGCCGCCANTCCGGAGTCGCCCTTC  645
Db  379 SerLeuGlyAlaAspValLeuProGluTyrLysLysLeuGlnIleAlaProGlnValTrp  398

Qy  646 ATCCTGTTGCACTGTGTGGGGGCACTGAGGAGCCACCTGGGATGGGCTTCATCCTGCTGCCACA  705
Db  399 ThrIleLeuHisTyrSerProPheLysAlaValTrpAspTrpLeuIleLeuLeuLeuVal  418

Qy  706 CTCTATGTGGCGTCACTGTGCCCTACAGCGTGTGTGTG------AGCACAGCACG  756
Db  419 IleTyrThrAlaValPheThrProTyrSerAlaAlaPheLeuLeuLysGluThrGluGlu  438

Qy  757 GAGCCCAGTGCC-------  
Db  439 GlyProProAlaThrGluCysGlyTyrAlaCysGlnProLeuAlaValAlaAspLeuIle  458

Qy  802 GTGGAGGTCCTCTTCATCCTTGACATGTGCTGAAATTTCCGTACCACATTCGTGTCCAAG  861
Db  459 ValAspIleMetPheIleLeuValAlaMetCysGlnProLeuAlaValAlaAspLeuIle  478

Qy  862 TCGGGCCCAGGTGGTGTTGCCCAAAGTACCATTGCCCACTACGTCACCACCTGGTTC  921
Db  479 AsnGluGluLeuValSerHisProGlyArgIleAlaAlaHisTyrPheLysGlyTrpPhe  498

Qy  922 CTGCTGGATGTCATCCGAGCCGTGCCCTTTGACCTGCTCACGCCCTACAAGGTGCAACGTG  981
Db  499 LeuIleLeuAspMetValAlaAlaIleIleProPheAspLeuLeu  511

Qy  982 TACTTCGGGGCC------------  
Db  512 IleProPheGlySerGlyLeuGluIleLeuIleGlyLeuLeuLysThrAlaArgLeuLeuArg  531

Qy 1024 CTGCTCGCGGCCTGCTTCCGCGGGCTGGACCTGGTGACCATCACAGCGCCGCCGTGGTGCTGACA  1083
Db  532 LeuValArgValAlaArgLysLeuAspArgTyrSerGluTyrGlyAlaAlaValLeuPhe  551

Qy 1084 CTGCTCCATGGCGGTGTTTCGGCCCTGCTTCCGCGGGCTGGTGTCGGCCCTGCCGTCTGGGTTTCACATT  1143
Db  552 LeuLeuMetCysThrPheAlaLeuIleAlaHisTrpLeuAlaCysIleTrpTyrAlaIle  571

Qy 1144 GGCCAGCGGGAGATCGAGAGCAGCAGAAATCCGAGCTGCCTGAGATTGGCGTGGCTGCGAGGAG  1203
Db  572 GlyAsnMetGluGlnProHisMetAspSer------ArgIleGlyTrpLeuHisAsn  588

Qy 1204 CTGGCCGCGCCGACTGGAGACTCCCTACTACCTGGTGCGGGCCGGAGGGCCGAGCTGGGAGGAGGAAC  1263
```

Thu May 8 08:49:44 2003   us-09-965-830-1_copy_6_3257.rsp   Page 4

```
Db   589 LeuGlyLysAspGlnIleGlyLysProTyr------------------------------------            |||||
                                                                                        :::::
Qy  1264 AGCCTCCGGCCCAGAGTGACAACTGACCAGCAGCCAGCGCAGGGCCAACGGACGGGGCTGGAG  1323
                                                                                        |||||||
Db   599 SerSerGly------                                                                 601
Qy  1324 CTGCTGGGGCGGGGGCCCCTGCTGCCCACCTCCCTTACTCTTCGGCACTTCAGC  1383
                                                                                        |||||
Db   602 ---LeuGlyProSerIleLeuLysProTyrTyrValThrAlaLeuThrPheSer  620
Qy  1384 AGCCTCACCACCGTGGGCCTTCGGCAATGGCGTCCGGCCAACACCGGAGAGATCTTC  1443
Db   621 SerLeuThrSerValGlyValProPheGlyAsnValSerProAsnThrAsnSerGluLysIlePhe  640
Qy  1444 TCCATCTGCACCATGTCATCGGGCCCGGTGGTGTTGTTTGGAACGTGACG  1503
                                                                                        :::|||::
Db   641 SerIleCysValMetLeuIleGlySerLeuMetTyrValSerIlePheGlyAsnValSer  660
Qy  1504 GCCATCATCCAACGCATGTACGGCCCCGGCGCTTTCTGTGTACCACAGCCCGCGCACTG  1563
                                                                                        |||||||
Db   661 AlaIleIleGlnArgLeuTyrSerGlyThrAlaArgTyrHisThrGlnMetLeuArgVal  680
Qy  1564 CGGCGACTACACCCGATCACCCAAGCCCATGCTGTGGAGTAC  1623
Db   681 ArgGluPheIleArgPheHisGlnIleProAsnProLeuArgGlnArgLeuGluTyr  700
Qy  1624 TTCCAGGCCACTGGGGTGAACAATGGCATCGACACCACCGAGCTGTGCTGCGAGAGCCTG  1683
Db   701 PheGlnHisAlaThrTrpSerTyrThrAsnGlyIleAspThrAlaValAlaValLeuLysIleGlyPhe  720
Qy  1684 CCTGACGAGCTGCGGCGAGACAATCGCCATTGCACCTGCACCTGCAGGTCCTGCAG---CTG  1740
Db   721 ProGluCysLeuGlnLysAlaAspIleCysLeuHisLeuLeuAsnArgSerLeuLeuGlnHisCys  740
Qy  1741 CCACTGTTTGAGGCGGGCGGCGCAGCCGGCGGCCGCTGTCTCTCGGCCCGGCCC  1800
Db   741 LysProPheArgGlyAlaThrLysGlyCysLeuArgAlaLeuLeuMetLysPheLysThr  760
Qy  1801 GCCTTCTGCACGGCGGAGACTACCTCCATCACCAAGGCAGGTGCCTCGCAG  1860
Db   761 ThrHisAlaProGlyTyrAspThrLeuValHisAlaGlyAspLeuLeuThrAlaLeuTyr  780
Qy  1861 TTTGTCTGTGCTCTGGCTCCATGGAGGTGCTCACGGTGGCAACCCTAGG  1920
Db   781 PheIleSerArgGlyTyrIleLeuIleLeuArgGlyAspValValAlaValAlaIleLeuGly  800
Qy  1921 AAGGGCGGACCTGATCGGGTGTGAGCTGCCCCGGGGGGCCAATGCC  1980
Db   801 LysAsnAspIlePheGlyLeuProLeuAsnLeuThrArgAlaArgProGlyLysSerAsnGly  820
Qy  1981 GACGTGAAGGGGCTGACGTACTCCGTCGTTGGCACCGTGGCCCTGCACGAC  2040
Db   821 AspValArgAlaLeuThrTyrCysAspLeuHisLysIleHisArgAspAspLeuLeuGlu  840
Qy  2041 AGCCTTGGCGTGTACCCCGAGTTTGCCCCGGAGTTTCCAGTGCTGGGAGCTC  2100
Db   841 ValLeuAspMetTyrProGluPheSerAspHisPheTrpSerSerLeu------  858
Qy  2101 AGCTACACCTG---------GGTGCTGGGGAGGCTCTCGCAGAGGTGGAC  2142
Db   859 ThrPheAsnLeuArgAspThrAsnMetIleProGlySerProGlySerThrArgGluLeuGlu  878
Qy  2143 ACCAGCTCCCTGAGGCGGGACAATACCCTTATGTCCACGCTGGACGACAGAT  2202
Db   879 GlyGlyPheSerArgGlnAsnArgLysSerPheArgArgThrAspProAsp  898
Qy  2203 GGGGAGCAGGGCCCCACGGTCTCCCAGCCCTGATGAGCCCCTGCTG  2262
Db   899 ThrGluGln------                                                                 901
Qy  2263 TCCCCTGGGTGCACCTCCTCATCCTGCCAAGCTGCTATCCCCAGGTCGAACAGCA  2322
```

```
Db   902 ---ProGlyGluValSerAla-------------------------LeuGlyProGlyArgAlaGly  914
Qy  2323 CCCCGGGCCTTCGTGCTTAGGTGCAGGTGGCACAGGGCCAGGCAGGGCAGGGCTTTGAAGGCTGAG  2382
Db   915 AlaGlyPro-----SerSerArgGlyArgProGlyGlyProTrpGluSerProSer  932
Qy  2383 GCTGGGCCCCTGTCTGGGGCACAGGGACCCCCA  2418
Db   933 SerGlyProSerSerArgSerGluSerSerProSerSer  952
Qy  2419 CTACGGCCTG-----CCCCCATGCCATGGAAATGTGCCCCAGGATCTTGAGCGCC  2466
Db   953 LeuArgLeuValProPheSerSerProArgProProGlyProGlyGluPro  972
Qy  2467 AGGGTAGTAGGAATGGCATTGAAGACGGGCGTGGGCTTCGGACCAGCCCAAG------  2514
Db   973 ---LeuMetGluAspCysCysAsnProLeuSerGlyValAlaPhe  991
Qy  2515 ----------------------------TTTCTCTTTC------CGGCTGTGGGGCCAGTGCT  2538
Db   992 SerGlyValSerAsnIlePheSerPheThrGlyAspSerArgArgGlyArgGlnTyrGlnGlu  1011
Qy  2539 GGCGCCGGAAGTGTAGGCAGCAGCAGCCCCTCCCGTGGCAGCGGCGGCCGTGCTCACTGTT---  2595
Db  1012 LeuProArgCys-----ProAlaProThrProSerLeuLeuAsnIlePro  1026
Qy  2596 ----------CCCCATGGGGCCCACCGCAAGGCCAAGGAACACAGAACACTG  2634
Db  1027 LeuSerSerProGlyArgArgProArgGlyLysAspValProValGluSerArg------Leu  1042
Qy  2635 GACAAGCTTCTGGGCACAGGGGTGACACAGACGCTGTCCTCAGAGGCAGGTG---CTGCAGATGCGGGGAA  2691
Db  1043 AspAlaLeuAlaGlnGlnAsnLeuAsnArgLeuGluThrArgLeuSerAlaAspMetAlaThr  1062
Qy  2692 GGACTGCAGTACACTTCGCCAGGCGTGTGCAGCTTGTCGCTCTTGCGCCCCCACAGGGAGGTCCG  2751
Db  1063 ValLeuGlnIleLeuLeuGln-ArgGlnMetThrLeu-----  1073
Qy  2752 TGCCCTCGGGCATCGGGACGAGGGGGCCGTGCCCAGCCACCTCCGGGCTTCTGCACCT  2811
Db  1074 ----------------------------ValProProAlaTyrSerAla-  1080
Qy  2812 CTGTGTGTGGGACACTGGGGGCGATCCTCCTACTGCCGCCCCCTACTGGCTCTGTCTTG  2871
Db  1081 ----ValThrThrProGlyProGlyProThrSerThrSerProLeuLeuProValSerPr  1099
Qy  2872 AGTGGGACTTGGGCCCCGCCACCCCGTCCGGGGGCGGCCTCCCCCATGGCACCCTGGCCCTGG  2931
Db  1100 oLeu-----ProThrLeuThrLeuAspSerLeuSerGln-----  1110
Qy  2932 GGTCCCCACGGTCTC------AGAGCTCCCCCACCTCGCAGCCCCAGCTTTCTGG  2982
Db  1111 -ValSerGlnPheMetAlaCysGluLeuGluProGly---AlaProGluLeuProGl  1129
Qy  2983 ACCTCCACCTCAGACTCAGAGAGCCCCGCCCTTCGACCCCACC  3042
Db  1129 nGlyGlyProThrArgArgLeuSerLeuProGlyGlnLeuLeuGlyAlaLeuThrSerGlnPr  1149
Qy  3043 CCTG  3046
Db  1149 oLeu  1150
```

RESULT 2
CIKE_DROME
ID   CIKE_DROME   STANDARD;   PRT; 1174 AA.
AC   Q02280;
DT   01-JUL-1993 (Rel. 26, Created)
DT   01-JUL-1993 (Rel. 26, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Potassium channel protein eag.
GN   EAG.
OS   Drosophila melanogaster (Fruit fly).
OC   Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;

```
OC   Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
OC   Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
OX   NCBI_TaxID=7227;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=91262635; PubMed=1840699;
RA   Warmke J., Drysdale R., Ganetzky B.;
RT   "A distinct potassium channel polypeptide encoded by the Drosophila
RT   eag locus.";
RL   Science 252:1560-1562(1991).
CC   -!- FUNCTION: PROTEIN EAG IS MOST PROBABLY A STRUCTURAL COMPONENT OF
CC       THE POTASSIUM CHANNEL AND MEDIATES THE POTASSIUM PERMEABILITY OF
CC       MEMBRANES.
CC   -!- SUBUNIT: DIMER (POTENTIAL).
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC   -!- MISCELLANEOUS: THE SEGMENT S4 IS PROBABLY THE VOLTAGE-SENSOR AND
CC       IS CHARACTERIZED BY A SERIES OF POSITIVELY CHARGED AMINO ACIDS AT
CC       EVERY THIRD POSITION.
CC   -!- MISCELLANEOUS: THE SEGMENT H5 IS THOUGHT TO LINE THE CHANNEL PORE.
CC   -!- SIMILARITY: TO THE MEMBERS OF THE POTASSIUM CHANNEL PROTEINS
CC       OF THE SH SUPERFAMILY.
CC   -!- SIMILARITY: CONTAINS 1 CYCLIC NUCLEOTIDE-BINDING DOMAIN.
CC   -!- SIMILARITY: CONTAINS 1 PAS (PER-ARNT-SIM) DIMERIZATION DOMAIN.
CC   -!- SIMILARITY: CONTAINS 1 PAS-ASSOCIATED C-TERMINAL (PAC) DOMAIN.
CC   ----------------------------------------------------------------------
CC       This SWISS-PROT entry is copyright. It is produced through a collaboration
CC       between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC       the European Bioinformatics Institute. There are no restrictions on its
CC       use by non-profit institutions as long as its content is in no way
CC       modified and this statement is not removed. Usage by and for commercial
CC       entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC       or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; M61157; AAA28495.1; -.
DR   PIR; A40853; A40853.
DR   FlyBase; FBgn0000535; eag.
DR   InterPro; IPR001622; K+channel_pore.
DR   InterPro; IPR000636; M+channel_nlg.
DR   InterPro; IPR001610; PAC.
DR   InterPro; IPR000700; PAS-assoc_C.
DR   InterPro; IPR000014; PAS_domain.
DR   InterPro; IPR000595; cNMP_binding.
DR   Pfam; PF00027; cNMP_binding; 1.
DR   Pfam; PF00520; ion_trans; 1.
DR   Pfam; PF00785; PAC; 1.
DR   SMART; SM00086; PAC; 1.
DR   SMART; SM00091; PAS; 1.
DR   SMART; SM00100; cNMP; 1.
DR   PROSITE; PS50042; CNMP_BINDING_3; 1.
DR   PROSITE; PS50112; PAS; 1.
DR   PROSITE; PS50113; PAC; 1.
KW   Ionic channel; Transmembrane; Ion transport; Voltage-gated channel;
KW   Glycoprotein; Phosphorylation; Repeat.
FT   TRANSMEM    227    246       SEGMENT S1 (POTENTIAL).
FT   TRANSMEM    269    291       SEGMENT S2 (POTENTIAL).
FT   TRANSMEM    314    335       SEGMENT S3 (POTENTIAL).
FT   TRANSMEM    343    369       SEGMENT S4 (POTENTIAL).
FT   TRANSMEM    372    393       SEGMENT S5 (POTENTIAL).
FT   TRANSMEM    442    467       SEGMENT H5 (POTENTIAL).
FT   TRANSMEM    471    493       SEGMENT S6 (POTENTIAL).
FT   DOMAIN       43     97       PAS.
FT   DOMAIN      113    165       PAC.
FT   NP_BIND     571    688       CNMP.
FT   CARBOHYD    262    262       N-LINKED (GLCNAC...) (POTENTIAL).
FT   CARBOHYD    412    412       N-LINKED (GLCNAC...) (POTENTIAL).
FT   CARBOHYD    424    424       N-LINKED (GLCNAC...) (POTENTIAL).
SQ   SEQUENCE   1174 AA;  126236 MW;  344C80DC06E4340E CRC64;
```

Alignment Scores:
Pred. No.:
Score:
Percent Similarity:
Best Local Similarity:

	2.38e-47
	1390.00
	47.57%
	31.55%

Length: 1174
Matches: 390
Conservative: 198
Mismatches: 398

US-09-965-830-1_COPY_6_3257 (1-3252) x CIKE_DROME (1-1174)

Query Match: 22.83%   Indels: 250
                        Gaps: 42

```
Qy       1 ATGGCG---GCCATGCGGGGCGGCCTCTGGGCCGTCGCCCCTCAGAACACCTTCCTGGACACCATCGCT 57
           |||||||   |||:||:|::||||::::|||||||||:::||||||||||||:::|||||:::|||||||
Db       1 MetProGlyGlyArgArgGlyLeuVal Ala ProGlnAsnThrPheLeuGluAsnIle Ile 20

Qy      58 ACGGCGCCTTCGACGGC--ACGCACAGTAACTTACGTTGCTGGGCAACCGCCAGGTGGCGGGG 114
           |||           :::   ||:|||||||||||||||:::||:|||||||||||||:|||||
Db      21 ArgArgSerArgAsnSerGlnProAspSerSerPheLeuGluLeuLeuAlaAsnAlaGlnIleLeuValAsp 40

Qy     115 CTCCTTCCCGTGGTCTACTGCCTCTAGTCTCATGGCGTTCTTCTGACCTCACGGGCTTCTCCCGGGCT 174
           |||:|||||||||:||||::|||::||:||||||:::|||:::||||||||||:::||||:|||||||
Db      41 PheProIleValTyrCysAsnGluSerPheCysLysIleSerGlyTyrAsnArgAla 59

Qy     175 GAGGTCATGCAGCGCGGCGGGGGTGT------GCCTGCGTCCTTCCGTTATGGCCGAGACACCAGT 228
           |||||||||||||||||||||||||      |||||||:::|||||:::||:::||||:||||||
Db      60 GluValMetGlnLysSerCysArgTyrValCysGlyLysPheMetThrTyrGlyGluLeuThrAsp 79

Qy     229 GAGGTCGTCCGCAACAGATCCCGCAAAGGCCCTGACGACGACAAGGAGTTCAAGGCTGAG 288
           |:::                    ||:::                      :::::
Db      80 LysGlyThrValGlyArgLeuGlyTyrThrLeuGluAsnGlnGlnAspGlnPheGlu 99

Qy     289 CTGATCCCGTACCGGAAGAGGCGGCTC----- 315
           :::||||::|||||||||||||||||
Db     100 IleLeuLeuTyrLysLysAsnLeuGlnCysGlyCysAlaLeuSerGlnPheGlyLys 119

Qy     316 ----------CCGTTCTGGTGTCTCCTGGATGTGACCATAAACGAATGAG 357
                     |||    |||     |||               ||||
Db     120 AlaGlnInThrGlnLeuThrProLeuTrpLeuLeuGlnValAlaProIleArgAsnGlu 139

Qy     358 AAAGGGGAGGTGGCTGCTCTTCTTCCTAGTCTCTCACAAGGACATCAGCGAAACCAGAACCGA 417
           :::
Db     140 ArgAspLeuValValLeuPheLeuLeuThrPheArgAspLleThrAlaLeuLeuLysGln--- 158

Qy     418 GGGGGCCCCGACGAGATGGAAGGAGACAGGTGGT------GGCGGGCGGCGATATGGCCGG 471
           |||              |||             |||      |||
Db     159 ----ProIleAspSerGluAspThrLysGlyValLeuGlyLeuSerLysPheAla Lys 176

Qy     472 --GCACGATCCAAAGGCTTCAATGCCAACCGGCGGGGGGCCGGAGCGGGCGGTGCTCTACCAC 528
           ||||                 |||                          |||
Db     177 LeuAlaArgSerVal--------------------ThrArgSerArg------------Gln 186

Qy     529 CTGTCCGGGCACCTGCAG-----------AAGCAGCCCAAGGGCAAGCAACAGCAAGCTCAATAAGGG 582
           :::||||::                   |||
Db     187 PheSerAlaHisLeuProThrLeuLysAspProThrLysGlnSerAsnLeuAlaHisMet 206

Qy     583 GTGTTGGGGAGAAACCAAACACTTGCCCTGAGTACAAAGTAGCGCCGCCATCCGGAAGTCGCC 642
           :::              ::||||:||:
Db     207 MetSerLeuSerAlaAspIleMetProGlnTyrArgGlnGluAlaProLysThrProPro 226

Qy     643 TTCATCCTGTTGCACTGTGGGGCCACTGAGAGCCCACTGGAGTGGCTTCATCCTGCTGCC 702
           |||||||||||||     ||||||||:
Db     227 HisIleLeuLeuHisTyrCysAlaThrGluSerProLeuGluTrpLeuIleLeuLeuCysLeu 246

Qy     703 ACACTCTATGTGGGCTGTCACTGTGCCCTACACGGTGTGTGTG------AGCACAGCACGG 756
           |||||||||||    |||   |||                    |||:::
Db     247 ThrPheTyrThrAlaLleMetValProTyrAsnValAlaPheLysAsnLysThrSerGlu 266

Qy     757 GAGCCCAGTGCGCCGCCGCCCGCGGGCCCCGCGGCGTCCTGTGGACCTCCGTGGAGTCCTC 816
           :::      |||||||||||||||||||||||||||:::||||||||:::||||||||||
Db     267 AspValSerProIleValAspVal IlePhe 281

Qy     817 ATCCTTGACATTGCTGAATTCCGTACCACACACTTCGTGTCCAAGTCGCGGCCCAGTGGTG 876
           ::|||||||||||||||||||||||||||||||||||:::||||:::|||||||||:||
Db     282 PheIleAspIleValLeuAsnPheHisThrThrPheValGlyProGlyGlyValVal 301

Qy     877 TTTGCCCAAAGTCCATTTGCCTCACACTACGTTCACCACCTTGGTTCGTCGTGGATTGTCATC 936
           |||       :::||||||||:::||||||::|||||:::||||||||:::|||||:::
Db     302 SerAspProLysValIleArgMetAsnTyrLeuLysSerIleProPheIleLeuAspLeuLeu 321
```

```
Qy    937  GCAGGCTGCCTGCCTTTGACCTTGCTACATGCCTTCAAG----------------GTCAACGTG          981
           |||   |||   |||   ||::|||:::|||                       :::
Db    322  SerCysLeuProTyrArgAspValPheAsnAlaPheAsnAlaPheAsnAlaAlaPheAsnArgAspGlyIleGlySer        341

Qy    982  TACTTCGGGGCCCATCTGCTGAAGACGGTGCGCCCTGCCGCCTGCTTCCG                          1041
           |||||||||||||||||||||||||||||||||||||||||||
Db    342  LeuPheSerAla------LeuLysValValArgLeuLeuArgLeuGlyArgValValAlaArg             359

Qy   1042  CGGCTGGGACCGGACGACTACAGCGACGACTACGGCCTGGGTTTC                               1101
           :::|||||||||||||||||||||||||::|||
Db    360  LysLeuLysArgThrTyrLeuGluThrArgGlyAlaAlaMetLeuIleLeuLeuLeuCysPheTyr          379

Qy   1102  GCCCTGCTCCGGCCACTGGGTCCTGCGTCCTGGTTTTTACAATTGGCCAGCGGGAGGATCGAG            1161
           |||||||||||||||||||||||||||::||||||||::
Db    380  MetLeuValAlaHisTrpLeuAlaCysHisIleTrpTyrSerIleGlyArgSerAspAlaAsp            399

Qy   1162  AGCAGCGAATCCGAGCTTGCCTGAGATTGGCTGGCCTCGAGGAGCTGGAG                          1221
           |||||||||||||||||||||||||::||||||||
Db    400  AsnGly--------IleGlnTyrSerTrpLeuTrpLysLeuAlaAlaAsnValThrGln                415

Qy   1222  ACTCCCTACTACCTGGTGTGGGCCCGGAGGCCAGCTGGAGGAACAGCTCCGGCCAGAGTGAC            1281
           |||||||||||||||||||||||::::|||||||
Db    416  SerProTyrSerTyrIleTrp----------------                                      422

Qy   1282  AACTGCGACGACGCAGAGGACGGGGCTGGAGCTGTGGGGGCCCGTCG                            1341
           :::|||||||||||||||||||||||||||||||
Db    423  ----------------SerAsnAspThrGlyProGluLeuValAsnGlyProSer                    435

Qy   1342  CTGCGGCCAGCGGCTACAACACCCCTACTTCGCCAGCCCTCACCAGGGTGGGC                      1401
           :::||||||||||||||||||||||||||||||||
Db    436  ArgLysSerMetTyrValThrAlaLeuTyrPheThrMetThrCysMetThrSerValGly               455

Qy   1402  TTCGGCAACGTGTCCGGCCAACACGGAGAGAGAGTCTTCCAATCTGCACCATGCTC                   1461
           |||||||||||||||||||||||||||||||||::
Db    456  PheGlyAsnValAlaIleAlaGluThrArgAspAsnGluLysLysValPheAsnThrIleAla            475

Qy   1462  ATCGGCGCGCCCTGAATGCAGCACGGCGGTGGTGTTTGGGAACGTGACGGCCATCATCAGGCCAATG        1521
           :::||||||||||||||||||||||||||||||||
Db    476  IleAlaAlaLeuLeuTyrAlaThrIlePheGlyHisValThrPheThrIleIleGlnIleMet            495

Qy   1522  TACGGCCCGCCGCGTTTCTGTACCACAGCCGACCTGCGGGACTACAACCGCATC                     1581
           ||||||||||||||||||||||||||::||||||||
Db    496  ThrSerAlaThrAlaLysTyrHisAlaSpMetLeuAsnAsnAlaValArgGluPheMetLysLeu          515

Qy   1582  CACCGTATCCCCAAGCCCCTCAAGCACGGGCCATGCTGGAGTACTTCCAGGCCACCTGGGCG            1641
           |||   |||||||||||||||||||||||||||||||
Db    516  HisGluValProLysAlaLeuSerGlyAsnValMetAspThrTyrValValSerThrTrpAla           535

Qy   1642  GTGAACAATGGCATCGACACCACCGAGCTGCTGCCAGAGACCTGCGGCTGCGCCA                     1701
           |||||||||||||||||||||||||||||||||||||
Db    536  MetThrLysGlyValLeuLeuAsnCysCysProLysAspMetLysAla                          555

Qy   1702  GACATCGGCCATGCAACCCTGCACCAAGGAGGTCCTG---CAGCTGCCACTGTTTGAGGCGGGC          1758
           ||||||||||||||||||||||||||||||   |||
Db    556  AspIleCysValHisLeuAsnArgLysValPheAspGluHisProThrPheArgLeuAla              575

Qy   1759  AGCCGGCCCTGCTCGGGCACGTCTCTGGCCCTGCGGGCACCGCCTTCGCACGCCGGGGC               1818
           :::|||||||||||||||||||||||||||||||
Db    576  SerArgAspGlyCysSerGlyThrArgLeuAlaLeuAlaMetHisSerHisSerAlaProGly           595

Qy   1819  GGAGTACCTCCATCCACCAAGGCGATGGCCCTCCAGGCCCTCTACTTTGTCTGCTCTGGGCTCC          1878
           |||||||||||||||||||||||||||||||||||::
Db    596  AspLeuLeuTyrHisThrGlyGluSerIleAspSerLeuCysSerPheIleLeuValThrGlySer        615

Qy   1879  ATGGAGGTGCTCAAGGGTGGCACCGTGTCGCACATCCTAGGGAAAGGGAGGCCGACCTGATCGGC        1938
           |||||||||||||||||||||||||::||||||||
Db    616  LeuGluValIleLeuAsnAspGlyValValAlaIleAlaLeuGlyLysGlyAspValPheGly           635

Qy   1939  TGTGAGCTGCCCCGGCGGGGAGCAGGTGGGTAAAGGGCTGACG                               1998
           |||   :::|||||||||||||||||||||||||
Db    636  AspGlnPheThrArgLysAspSerAlaLeuValGlyLnSerAlaLeuAlaAsnValArgAlaLeuThr       655

Qy   1999  TACTCGCGTCCTGCAGTGTGTCTGCACACAGCAGCCTTCGCGGTGTACCCC                        2058
```

```
Db    656  TyrCysAspLeuHisIleAlaIleLysLysArgAspLysLysGluLeuLeuGluValLeuAspPheTyrSer    675
           |||||||
Qy   2059  GAGTTTGCCCCGGCGGCCTTCAGTCGGTCGGTCCCTCCGAGGGAGAGTCTCAGCTACACAGCCCTGGGCT     2118
           :::|||   |||
Db    676  AlaPheAlaAsnSerPheAlaArgAsnLeu------ValLeuThrTyrAsnLeu----            691

Qy   2119  GGGGGAGGCCTCTGCGAGAGGTGGACACCAGCTCCCTGACGGGCGGACAATACCCTTATGTCC          2178
           |||||||||||||||||||||||||||||||||||::
Db    692  ------------------------------------ArgHisArg                             694

Qy   2179  ACGGTGGAGGAGAAGAGGAAGAGGACAGATGGGGAGGCAGGGCACGGGTCCCCACAGCCAGCT          2238
           :::|||   |||
Db    695  LeuIleLeuPheArgArgValAlaAlaAspValLysSerGluLeuAlaGluArgArgLys              714

Qy   2239  GATGAGCCCTCACCAGCCCCTGCTGTCCCCGGCTGCACCTCCTCATCCCTCAGCTGCCAAG             2298
           ||||||||||||||||||||||||||||||||::
Db    715  AsnGluProGlnLeuProGln------AsnGlnAspHisLeuValArgLys                        729

Qy   2299  CTGCTATCCCCCAGCTCGAACAGCACACCCCGGCCGCCCGTCCGTCGAGAGGAGGAGGAGGCC           2358
           |||   |||:::|||
Db    730  IlePheSerLysPheArgArgThrProGlnValGlnAlaGlySerLysGluLeuLeuGly               749

Qy   2359  AGGGCAGGGGGCTTTGAAGGCTGAGGGCTGGC                                           2388
           |||||||||||||||::|||||||||||
Db    750  GlySerGlyGlnSerAspValGlyLysGlyGluValGluArgGluProSerArgAspSerArgArgValVal   769

Qy   2389  CCCTTCGCTCCCCCCACGGCGGCCCGCGACCCCAGCCCAAGGTTCTTCTTTCCGACGAGGAGC           2442
           |||   |||
Db    770  ProLysAlaProGlyLeuGlnLeuAlaSerGlnAsnAlaThrLeuAlaArgGlnAlaThrIleAsp        789

Qy   2443  -------------------AAATGTGCCCCAGATCGAGCCCCCAGGGTAGTAGAGTAGCATT           2484
           :::||||||||||||||||||||||||
Db    790  GluGlyGlyGluValAspSerSerProSerArgAspSerAspSerArgArgValVal----Ile          807

Qy   2485  GAAGACGGCTGTGGCTCGCGGACCCAGCCCAAGTGTCTTCTTTCGACGAGGAGCTGTCGGCCG           2544
           |||||||||||||||||||||||||::||||||||
Db    808  GluGlyValAlaAlaValSer----------------SerAlaThrValGlyProSerProPro         822

Qy   2545  GAAATGTAGCAGCAGCCCC-------------------------TCCCCTGGACCAGAGAGGCGGC          2583
           |||||||||||||||||||
Db    823  ValAlaThrThrSerAlaAlaAlaAlaGlyValSerGlyGlyProGlySerGly                    842

Qy   2584  -----------------CTGCTCACT                                               2592
           :::|||||||
Db    843  GlyThrArgValValAlaIleValThrLysSerAlaAspArgAsnLeuAlaLeuLeuArgGluArg        862

Qy   2593  ---GTTCCCCATGGCCCCCACGCAGCCCCAGCGCAAGGAAACACACACTGACAAG---CTTCGG          2646
           :::|||
Db    863  GlnIleGluMetAlaSerSerArgAlaThrThrSerArgAspThrTyrAspThrThrGlyLeuArg        882

Qy   2647  CAGGCGGTGACGAGAGCTGTCAGGAGCAG----------GTGCGTGCAGATG-----              2685
           ::::|||::::|||
Db    883  GluThrProProThrLeuGlnArgAspLeuValAlaThrAlaThrArgValLeuAspMetLysVal        902

Qy   2686  ------CGGGAAGGACTGCAGTCACTTCGGCCAG-------GCTCGTGCAGCCTTTGTCCTG            2730
           |||||||||||||||||||::::|||
Db    903  AspValArgArgLeuGluGlnArgMetGlnAsnArgIleGluGlyArgIleGluGluLeuArg           922

Qy   2731  GCGCCCCCACAGGGAGGGGTCCCGTGCCCCTCGCCGGAGAGGGGCGCCGCCCAGC                    2790
           |||||||||||||||||||||||||||||||||||::
Db    923  GlyGlyLeuValLysArgLeuValLeuAlaSerSerGlyGlyAsnAlaAlaProAlaAspAsn          942

Qy   2791  ACCTTCGGGGCGTTCTGCAGCCT----------CTGTGT-----                            2817
           |||:::|||
Db    943  SerSerGlyGlnThrThrProGlyLysAspGluIleCysSerAlaGlyGlyAlaGlyGlyGly          962

Qy   2818  ------------GTGGACACT---                                                 2826
           |||||||||||
Db    963  GlyThrProThrArgGlnAlaProProThrSerAlaValThrSerProValThrVal                982
```

```
Db    983 IleThrIleSerSerProGlyAlaSerGlySerGlyThrGlyAlaGlyAlaGlyAlaGly 1002
Qy   2827 ---------------GGGGCATCCTCCTACTGCCTCAGCCCCCAGCTGGGCTCTGTCTTGAGT 2874
Db   1003 SerAlaValAlaAlaGlyAlaGlyAlaGlyLeuLeuAspProGlyAlaThrValSer 1022

Qy   2875 GGGGACTTGGCCCACCCTCGTCCGGGGCGCTCCTCCCCCATG-------- 2916
Db   1023 SerAlaGlyGlyAsn--------GlyLeuGlyProLeuMetLeuLeuLysLysArgArgSer 1039

Qy   2917 -----GCACCCTGGCCCTGGGCGTCCCCAGCGTCTCAGAGCTCCCCCTGGCCT 2964
Db   1040 LysSerGlyLysAlaAlaProAlaPro------ProGluGlnThrLeuAlaSerThrAlaGly 1057

Qy   2965 CGAGCCACAGCT-----------
Db   1058 ThrAlaThrAlaAlaProAlaGlyValAlaGlySerGlyMetThrSerSerAlaProAla 1077

Qy   3013 TCAGGAGACCTCTGCTCTGAGCCCAGCACCCCCTCCTCCTTCTGAGGAGG 3072
Db   1078 SerAlaAspGlnGlnGlnGlnHisGlnSerAlaAlaAspGlnSerProThrHisProGly 1097

Qy   3073 GCTAGGACTGGGCCCGCAGAG--------------
Db   1098 AlaGlyLeuLeuHisLeuArgLeuLeuGluAspPheThrHisAlaGlnLeuProSer 1117

Qy   3115 ACCAGCACTGGAGAGCCC----CCACCAGGGTCAGGGGGGCCTGGCCTTGCCCTTGGGAC 3168
Db   1118 ThrSerSerGlyGlyAlaGlyGlyGlyGlyGlySerGlySerGlyAlaThrProThrThr 1137

Qy   3169 CCCCACAGCCTGGAGATGGGTCGTTATTGGCCTCCCCATGGCTCTCTGGCCACA 3216
Db   1138 Pro----------ProProThrIleAlaGlyGlyGlySerGlySerGlyThr 1150
```

RESULT 3
CNG3_BOVIN

```
ID   CNG3_BOVIN      STANDARD;      PRT;   706 AA.
AC   Q29441;
DT   01-NOV-1997 (Rel. 35, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Cyclic-nucleotide-gated cation channel alpha 3 (CNG channel alpha 3)
DE   (CNG-3) (CNG3) (Cyclic nucleotide gated channel alpha 3) (Cone
DE   photoreceptor cGMP-gated channel alpha subunit).
GN   CNGA3 OR CNCG3.
OS   Bos taurus (Bovine).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
OC   Bovidae; Bovinae; Bos.
OX   NCBI_TaxID=9913;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Testis;
RX   MEDLINE=94211295; PubMed=7512693;
RA   Weyand I., Godde M., Frings S., Weiner J., Mueller F., Altenhofen W.,
RA   Hatt H., Kaupp U.B.;
RT   "Cloning and functional expression of a cyclic-nucleotide-gated
RT   channel from mammalian sperm.";
RL   Nature 368:859-863(1994).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Kidney;
RX   MEDLINE=94224768; PubMed=8170936;
RA   Biel M., Zong X., Distler M., Bosse E., Klugbauer N., Murakami M.,
RA   Flockerzi V., Hofmann F.;
RT   "Another member of the cyclic nucleotide-gated channel family,
RT   expressed in testis, kidney, and heart.";
RL   Proc. Natl. Acad. Sci. U.S.A. 91:3505-3509(1994).
CC   -!- FUNCTION: COULD BE RESPONSIBLE FOR CGMP-INDUCED CALCIUM ENTRY IN
CC       CELLS OTHER THAN SENSORY CELLS. MIGHT BE INVOLVED IN CHEMOTAXIS OF
CC       SPERM.
CC   -!- SUBUNIT: FORMS FUNCTIONAL HETEROOLIGOMERIC CHANNELS WITH CNG4 IN
CC       VITRO.
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC   -!- TISSUE SPECIFICITY: TESTIS, KIDNEY, RETINAL CONE, AND HEART.
CC   -!- SIMILARITY: BELONGS TO THE CYCLIC NUCLEOTIDE-GATED CATION CHANNEL
CC       FAMILY.
```

```
DR   EMBL; X89600; CAA61759.1; -.
DR   EMBL; X76485; CAA54023.1; -.
DR   InterPro; IPR000636; M+channel_nlg.
DR   InterPro; IPR000595; cNMP_binding.
DR   Pfam; PF00027; cNMP_binding; 1.
DR   Pfam; PF00520; Ion_trans; 1.
DR   SMART; SM00100; cNMP; 1.
DR   PROSITE; PS00888; CNMP_BINDING_1; 1.
DR   PROSITE; PS00889; CNMP_BINDING_2; 1.
DR   PROSITE; PS50042; CNMP_BINDING_3; 1.
KW   Ionic channel; Ion transport; cAMP-binding; Transmembrane;
KW   Multigene family; Vision.
FT   TRANSMEM    186    206    POTENTIAL.
FT   TRANSMEM    322    342    POTENTIAL.
FT   TRANSMEM    398    418    POTENTIAL.
FT   NP_BIND     501    624    CAMP.
FT   BINDING     568    568    CAMP (POTENTIAL).
FT   BINDING     583    583    CAMP (POTENTIAL).
SQ   SEQUENCE    706 AA;   81132 MW;   F490DCD29B56239 CRC64;
```

Alignment Scores:

Pred. No.:		
Score:	1.33e-10	Length: 706
Percent Similarity:	439.00	Matches: 145
Best Local Similarity:	42.56%	Conservative: 101
Query Match:	25.09%	Mismatches: 208
DB:	7.21%	Indels: 124
	1	Gaps: 19

US-09-965-830-1_copy_6_3257 (1-3252) x CNG3_BOVIN (1-706)

```
Qy   424 CCGGACAGATGGAAGGAGACAAGCTCAATAAAGGGGTTGTTTGGGGAGAAACCAAAC 483
Db   127 ProAspArg-------------------------GlyArgSerAlaTrpProLeuAlaArgAsnAsn 140

Qy   484 GGCTTCAATGCCAACCGGCGGAGCCGGAGCCGGCGGCGGCGCGTGCTCTACCACCTGTCCGGCCACCTG 543
Db   141 ThrAsnThrCysAsnAsnSerGlyLys----------------------AspAspLysAla 153

Qy   544 CAGAAGCAGCCCAAGGGGCAAGCAACACAAGGCTCAATAAAGAGGGTGGTGGCCGGCGGCACGATCCAAA 603
Db   154 LysLysGluGluGluGluLysLysGluGluLysLys-----------------GluAsnProLys 169

Qy   604 TTGCCCGAGTACAAAGTAGCCCGCCCATCCGGAAAGTCGCCCTTCATCCTGTTGCACTGTGG 663
Db   170 LysGluGluLeuLysLysLysAspSerSerValValMetAspProSerSerAsnMetTyrTyrHis 189

Qy   664 GCACTGAGAGCCACCTGGGATGGCTTCATCCTGCCCACACTCTATGTGGCTGTCACT 723
Db   190 TrpLeuThrVal--------------------------IleAlaValPro 197

Qy   724 GTGCCCTACCAGCGTGGTGGTGAGCACGACGGGAGCCCAGTGCCGCGGCGCCCCGCCC 783
Db   198 ValPheTyrAsnTrpCysLeuLeuValCysArg-------------------- 208

Qy   784 AGCGTCTGT--------------
Db   209 ---AlaCysPheAspGluLeuGlnSerGlnHisLeuMetLeuTrpLeuValLeuAspPheTyr 227

Qy   799 GCCGTGGAGGTCCTCTCATCCTTGACATTGTCTGAATTCCGTACCACATTCGTGTCC 858
```

```
Db  228 SerAlaAspIleLeuTyrGlyMetAspMetLeuValArgAlaArgThrGlyPheLeu---  246
        :::||                     |||
Qy  859 AAGTCGGGCCCAGGTGGTGTTGGCCCAAAGTCCATTGCCTCACTACGTC---ACCACC  915

Db  247 GluGlnGlyLeuMetValMetAspAlaSerArgLeuTrpLysHisTyrThrGlnThrLeu  266
                |:::                    |||
Qy  916 TGGTTCCTTGCTCGGATGTCATTCGACGGCTGCCCCTTTTGACCTGCTACATGCCTTCAAGGTC  975

Db  267 HisPheLysLeuAspValLeuSerLeuValProThrAspLeuValTyr---PheLysLeu  285
        ||||||:::|||                  |||
Qy  976 AACGTGTACTTCCGGGGCCCCATCTGTCTGAAGACCGGTGCGGCCTGCTGCTGCGGGCTG  1035

Db  286 GlyMetAsnTyrProGlu------LeuArgPheAsnArgLeuLeuLysLeuAlaArgLeu  303
        ||||||                |||||||::|||
Qy 1036 CTTCCGCGCGCTGACCGGGTACTCG-------------------CAGTACAGC  1068

Db  304 PheGluPhePheAspArgThrGluThrArgThrAsnTyrProAsnMetPheArgIleGly  323
                                        :::
Qy 1069 GCCGTGGTCGCTGACACTGCTCCTGCCGGTGTTCCCCTCGCCACTGGGTCGCCTGC  1128

Db  324 AsnLeuValLeuTyrIleLeu---IleIleIleHisTrpAsnAlaCys  338
                            :::||||              |||
Qy 1129 GTCTCGTTTTTACATTGGCCAGCGGGGAGATCGAGAGCAGGAATCCGAGCTTGCCTGAGATT  1188

Db  339 IleTyrPheAlaIleSerLys---------PheIle  347
        :::|| :||                |||
Qy 1189 GGCTGGCTGCCAGGGAGCTGGGCCGCCGCGCACTGGACTCCCCTACTACCTGGGGGCCCGGAGG  1248

Db  348 GlyPhe-----------  349
Qy 1249 CCAGCTGGAGGGAACAGCTCCGGCGCCAGAGTCACGACACTGCACGCAGCAGGCCAAC  1308

Db  350 --------GlyThrAspSerTrpValTyrProAsnValSerAsnProGluTyrGlyArg  366
Qy 1309 GGGACGGGGCTGGGAGCTGCTGGGGGCGCCCGGTGCCTGCGGGAGCCCTACACTCACCTCCCTC  1368

Db  367 ---------LeuSerArgLysTyrIleTyrSerLeu  375
Qy 1369 TACTTCGCACTCAGCAGCCTCACCACGCGTGGGGCTTCGGCCAACGCAGGGAC  1428

Db  376 TyrTrpSerThrLeuThrLeuThrThrIleGly---GluThrProProProValLysAsp  394
Qy 1429 ACCGAGAAGATCTTCTCCCATCTGCACCATGCCATCGGACGATGCACGGTGGTG  1488

Db  395 GluGluTyrLeuPheValValIleLeuAspPheLeuValGlyValLeuIlePheAlaThrIle  414
Qy 1489 TTTGGGAACGTGACGGCCATCANCCACCGCATGTACGGCGCTTTCTGTGTACCACAGC  1548

Db  415 ValGlyAsnValGlySerMetIleSerAsnMetAsnAlaSerArgAlaGluPheGlnAla  434
Qy 1549 CGCACGCCGCGACCTGCGGCGACTACATCCGCATCCACGTATCCCCAAGCCCCTCAAGCAG  1608

Db  435 LysIleAspSerIleLysSerGlnTyrMetGlnPheArgLysSerValPheArgAspGluThr  454
Qy 1609 CGCCATGCTCGGAGTACTTCCAGGCGCCACCGGGGTGAACAATGCACCTGCACAAGGAG  1668

Db  455 ArgValIleArgTrpPheAspTyrLeuTrpAlaAsnLysLysThrValAspGluLysGlu  474
Qy 1669 CTGCTCGAGAGACCTCCCTGACGAGCTGGGGCCGAGACANCGACCTGCACAAGGAG  1728

Db  475 ValLeuLysSerLeuProAspLysLeuLysAlaGluIleLeuAsnValHisLeuAsp  494
Qy 1729 GTCCTG---CAGTCGCCACTGTTTGAAGGGCGGCCAGCCGGGCTCCTGGGGCAACTGTCT  1785

Db  495 ThrLeuArgLysValArgIlePheGlnAspCysGluAlaGlyLeuLeuValGluLeuVal  514
Qy 1786 CTGGCCCTGCGGCCCGGCCGTTCTGCCAGGCGGGAGTACCTCATCCACCAAGGCGGATGCC  1845

Db  515 LeuLysLeuArgProAlaValPheSerProGlyAspTyrIleCysLysLysGlyAspIle  534
Qy 1846 CTGCAGGCCCTCTACTTTGTCTGTGCTGGCTCCTTGGAGGTGCTC---AAGGGTGGC  1899

Db  535 GlyArgGluMetTyrIleIleLysGlyLysLeuAlaValValAlaGlyAlaGluGlyIle  554


Qy 1900 ACCGTGCTCGCCATCCTAGGGAGGGACGGACCTGATCGGC-------------TGTGAGCTG  1947
        |||            :::|||||||||                          :::
Db  555 ThrGlnPheValValLeuGlyAspGlySerTyrPheGlyIleIleSerIleLeuAsnIle  574

Qy 1948 CCCCGGCGGGGAGCAGGTGGTAAAGGCCAATGCCGACGTGAAAGGGCTGAGGTACTCGGTC  2007
        |||                                      |||
Db  575 LysGlySerLysSerGlyAsnArgArgThrAlaAsnIleArgSerIleGlyTyrSerAsp  594

Qy 2008 CTGCACGTGTCTGCAGCTGGCTGGGCTGCCACCAAGACGCCTTGGCGTCGAGGTACCCGAG  2061
        :::|||||                                  |||
Db  595 LeuPheCysLeuSerLysLysAspAspLeuMetGluAlaLeuMetThrGluTyrProGlu  612
```

RESULT 4

```
CNG1_RAT
ID   CNG1_RAT      STANDARD;      PRT;    683 AA.
AC   Q62927; O08659;
DT   01-NOV-1997 (Rel. 35, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   cGMP-gated cation channel alpha 1 (CNG channel alpha 1) (CNG-1) (CNG1)
DE   (Cyclic nucleotide gated channel alpha 1) (Cyclic nucleotide-gated cation channel 1)
DE   (Rod photoreceptor cGMP-gated channel alpha subunit).
GN   CNGA1 OR CNCG1 OR CNCG.
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Sprague-Dawley;
RX   MEDLINE=97197878; PubMed=9045728;
RA   Bradley J., Zhang Y., Bakin R., Lester H.A., Ronnett G.V., Zinn K.;
RT   "Functional expression of the heteromeric 'olfactory' cyclic
RT   nucleotide-gated channel in the hippocampus: a potential effector of
RT   synaptic plasticity in brain neurons.";
RL   J. Neurosci. 17:1993-2005(1997).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Sprague-Dawley;
RX   MEDLINE=9287732; PubMed=9142860;
RA   Ding C., Potter E.D., Qiu W., Coon S.L., Levine M.A., Guggino S.E.;
RT   "Cloning and widespread distribution of the rat rod-type cyclic
RT   nucleotide-gated cation channel.";
RL   Am. J. Physiol. 272:C1335-C1344(1997).
RN   [3]
RP   SEQUENCE OF 521-683 FROM N.A.
CC   -!- FUNCTION: VISUAL SIGNAL TRANSDUCTION IS MEDIATED BY A G-PROTEIN
CC       COUPLED CASCADE USING CGMP AS SECOND MESSENGER. THIS PROTEIN CAN
CC       BE ACTIVATED BY CYCLIC GMP WHICH LEADS TO A OPENING OF THE CATION
CC       CHANNEL AND THEREBY CAUSING A DEPOLARIZATION OF ROD
CC       PHOTORECEPTORS.
CC   -!- SUBUNIT: HOMOTETRAMER OR HIGHER OLIGOMER.
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC   -!- TISSUE SPECIFICITY: ROD CELLS IN THE RETINA.
CC   -!- SIMILARITY: BELONGS TO THE CYCLIC NUCLEOTIDE-GATED CATION CHANNEL
CC       FAMILY.
```

```
DR   EMBL; U48803; AAA92110.1; -.
DR   EMBL; U93851; AAC53139.1; -.
DR   EMBL; U76220; AAC17594.1; -.
DR   InterPro; IPR000636; M+channel_nlg.
```

```
DR   InterPro; IPR000595; cNMP_binding.
DR   Pfam; PF00027; cNMP_binding; 1.
DR   Pfam; PF00520; Ion_trans; 1.
DR   SMART; SM00100; CNMP; 1.
DR   PROSITE; PS00888; CNMP_BINDING_1; 1.
DR   PROSITE; PS00889; CNMP_BINDING_2; 1.
DR   PROSITE; PS50042; CNMP_BINDING_3; 1.
KW   Ionic channel; Ion transport; cAMP-binding; Transmembrane;
KW   Multigene family; Vision.
FT   DOMAIN        1    155       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    156    176       H1 (POTENTIAL).
FT   DOMAIN      177    189       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    190    208       H2 (POTENTIAL).
FT   DOMAIN      209    232       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    233    252       H3 (POTENTIAL).
FT   DOMAIN      253    290       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    291    313       H4 (POTENTIAL).
FT   DOMAIN      314    365       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    366    385       H5 (POTENTIAL).
FT   DOMAIN      386    469       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    470    490       H6 (POTENTIAL).
FT   DOMAIN      491    683       CYTOPLASMIC (POTENTIAL).
FT   NP_BIND     478    600       CGMP (POTENTIAL).
FT   BINDING     537    537       CGMP (POTENTIAL).
FT   BINDING     552    552       CGMP (POTENTIAL).
FT   CARBOHYD    416    416       N-LINKED (GLCNAC...) (POTENTIAL).
FT   CONFLICT      3      3       N -> KV (IN REF. 2).
FT   CONFLICT    141    141       E -> G (IN REF. 2).
FT   CONFLICT    224    224       E -> K (IN REF. 2).
SQ   SEQUENCE    683 AA;  79228 MW;  62D49F64F8D917C5 CRC64;

Alignment No.:
Pred. No.:
Score:                   3.11e-10
Score:                   429.50
Percent Similarity:      41.73%
Best Local Similarity:   22.59%
Query Match:             7.05%
DB:

Length:          683
Matches:         157
Conservative:    133
Mismatches:      278
Indels:          127
Gaps:             22

US-09-965-830-1_copy_6_3257 (1-3252) x CNG1_RAT (1-683)

Qy   181  ATGGAGCGGGGCGGGTCGCTCTCCTTCCTTCCTTTATGGGCACCAGCACCTGCAGCTCCGGC  240
                ||||| ||||||||||||||
Db    30  MetGluAsnGly---AlaCysSerSerPheSerAspAsnAspAsnGlySerLeuSerGlu     48

Qy   241  CAACAGATCCGCAAGGCCCTGGACGAGCACAAGAGAGTTCAAGGCTGAGCTGATCCTGTAC  300
               :::                    :::||
Db    49  Glu------------SerGluAsnGluAspSerLeuPheArgSerAsn-----SerTyr     62

Qy   301  CGGAAGAGACGGGGCGGGTCCCGTTCGTGTGTCTTCGGTGATGATACCCATAAAGAAATGAG---  357
               |||  |||
Db    63  ArgArgArgGly----------------------------ProSerGlnArgGluHis      72

Qy   358  ------AAAGGGGAGGTGGCTCGTCTTCTCAGTGTCTCACAAGGACATACGCAGAAACCAAG  411
                ::
Db    73  TyrLeuProGlyThrMetAlaLeuPheAsnValAsnAsn--------SerSerAsnLys     89

Qy   412  AACGGAGGGGGCCCGACAGATGGAAGGAGAGACAGTGGTGGCCGGGTCCGGCCGGATATGGGCCGG  471
               :::::
Db    90  AspGlnAspProLysGluLysLysLys----------------------------Lys      99

Qy   472  GCACGATCCAAAGGCTTCAAATGCCAACCGGCGGGCGGGAGCGGGCCGTGCTTCTCCACACTG  531
               |||  :::::
Db   100  LysLysGluLysLysSerLysAlaAspLysAspLysLysGluSer---------------    113

Qy   532  TCCGGGCACCTGCAGAAGCCCAAGGGCCAAGCACACAAGCTCAATAAGGAGGGGTGTTTGGG  591
               :::::
Db   114  ------LysLysAspProGlyLysGlyLysLysLys------------------------    129

Qy   592  GAGAAACCAAACTTGCCTGAGTACAAAGTAGCCGCCCACCGGAAGTCCCCTTCATCCTG   651
               ::: |||  ::::
Db   130  LysLysGlyGluGluProLysGlyGlyGlyGluGlyLys                        149

Qy   652  TTGCACTGTGGGGCCACTGAGAGCCACCTGGGATGGGGCTTCATCCTGCCTTCGGCCACACTCTAT  711
               |||  :::  :::
Db   150  AspProSerGlyAsnMetThrTyrTyrAsnTrpLeuPheCysIleThrLeuProValMetTyr   169

Qy   712  GTGGCTGTCACTGTGCCTTACCACACATTCTGTGCCTTACACGCGTGTGTGTGAGCACACGACGGGGAGCCCAGTGCCGCC  771
               :::
Db   170  AsnTrpThrMetIleIleAlaArgAlaCysPheAspGluLeuGlnSerAspTyrLeuGlu    189

Qy   772  CGGGGACCGCCCACGGTCTGTGACCTGGCCGGTGGAGGGTCCTCTCTTCATCCTTGACATTGTG  831
               ::: ::::::::
Db   190  TyrTrp-----LeuIlePheAspTyrValSerAspTrpIleTyrLeuAlaAspMetPhe     207

Qy   832  CTGGAATTTCCGTACCACACATTCGTGTCCTTCCAAGTCGGGCCAGTGTTCCCCAAAGTCC  891
               ||||| :::::
Db   208  ValArgThrArgThrGlyTyrLeuGluGlnGlyLeuLeuValLysAspGluLeuLysLeu    227

Qy   892  ATTTTGCTTCCTCCACTACAGTCACCACCTGGTTCCTGCCTGGGATGTCATCGCAGGGCTTGCCCCTTT  951
               |||
Db   228  IleGlyLysTyrLeuSerAlaAsnLeuGlnPheLysLeuAsnProThr               247

Qy   952  GACCTGCTACATCGCCTTCAAAGGTCAAACGTGTACTTCGGG---------  996
               |||||||||||:::  ||||||
Db   248  AspLeuLeuTyr----PheLys---------PheGlyTrpAsnTyrProGlyIleArg      262

Qy   997  CTGCTGAAGACGCGGTTGCGGCGCGTCTTGCAGTCCGGGCTTGGCAGCCGGTACTCCGGGTAC  1056
               |||  :::||:::
Db   263  LeuAsnArgLeuLeuArgIleSerArgMetPheGluPhePheGlnArgThrGluThrArg    282

Qy  1057  TCGGAGTACAGCGACGGGCGTGGTG---CTGACACTGCTCATGGCCGTGTTCGCCCTGGTCCG  1113
               ||||||||||| :::
Db   283  ThrAsnTyrProAsnIleIlePheArgIleSerAsnLeuValMetTyrIleValIleIleIleIle   302

Qy  1114  CACTGGGGTCGCCTGCCTGCGTGGTTGTGGTTTTTACATTGGCCAGCGGGGAGATCGAGAGCAGCAGGAATCC  1173
               ||||||
Db   303  HisTrpAsnAlaCysValTyrTyrSerIleSerLysLysAlaIleGluGlyPheGlyAsnAspThr   322

Qy  1174  GAGCTG------CCTGAGATTGGCTGGCTGCGAGGACTGGCCGGCGA  1215
               ::: |||
Db   323  ArgLeuAlaArgLys                                  339

Qy  1216  CTGGAGACTCCCTACTACCTGGTGGGGCCGGGAGCCCAGCTGGGAGGGAACAGCTCCGGCCAG  1275
               :::
Db   339  ------------------------------------------------------------   339

Qy  1276  AGTGACAACTGCGAGCAGCAGGAGCGCAACGGAGGCGTGGAGCTGCTGGGCGGC  1335
               :::
Db   339  ------------------------------------------------------   339

Qy  1336  CCGTCGCTGCCGCAGGCGCCTACAACATCACCTCCCTTCGCACTTCACAGCGTCACCAGC  1395
               ||| :::||||||| :::
Db   340  --------------------TyrValTyrSerLeuTyrTrpSerThrLeuThrThr       353

Qy  1396  GTGGGCGCTTCGGCAACGTTGCCGCACACAGACGAGATCTTCGCACCATCGGACC  1455
               :::|||
Db   354  IleGly----GluThrProProProValLeuAspSerGluTyrValPheValValAlaAsp   372

Qy  1456  ATGCTCATCGGGGCCCTGATGCACGCGGTGGTGTTTGGGAACGTGACGGCCATCATCCAG  1515
               ||||||||| :::||:::
Db   373  PheLeuIleGlyValLeuIlePheAlaThrIleValGlyAsnIleGlySerMetIleSer    392

Qy  1516  CGCATGCTACGGCGCCGCGCCGCTTCTTGTGTACCACAGCACCGCCCGACCTGCACACATC  1575
               |||:::||||||||||||||
Db   393  AsnMetAsnAlaAlaArgAlaGluPheGlnSerArgMetAspAlaIleLysGlnTyrMet    412

Qy  1576  CGCATCCACGTATCCCAAGCCCCTCAAGGAGCAGCAGCGCATGCGGGGCACGTTCAGGCACC  1635
               :::|||  ||||||
Db   413  PheLeuLeuValSerLysGlyMetGlyLysGluSerTrpProAspTyrLeu           432

Qy  1636  TGGGGCGGTGGACAAATGGCATGGCAGCACACCAGAGCTGCAGAGCCTCCCGGACGAGCTG  1695
               |||
Db   433  TrpThrAsnLysLysThrValAspGluArgGluValLeuLysArgTyrLeuProLysGlyLeu  452

Qy  1696  CGGCGAGACATCGCCACCTGCACAAGGAGGTCCTG---CAGCTGCCACTGTTTGAG  1752
```

```
Db   453 ArgAlaGlyIleAlaIleAsnValHisLeuAspThrLeuLysGlysValArgIlePheAla 472
          ||| ||||||::|||:::::||||||||||:::||||  :||||    |||::||||||
Qy  1753 GCGGGCCAGCCGGCGCTGCCTGCGGGCACTGTCTCTGGGCGCCGCTTCTGCACG 1812

Db   473 AspCysGluAlaGlyLeuLeuValGluLeuValLeuLysLeuGlnProGlnValTyrSer 492
          |||   ||| |||:||||||||| |||::|||||::|||||||:::|||  :::||||:||
Qy  1813 CCGGGCGGAGTACCTCATCCACCACCAAGGCGATGCCCTGCAGGCCCTCTCACTTTGTCTGTCT 1872

Db   493 ProGlyAspTyrIleCysLysLysGlyAspIleGlyArgGluMetTyrIleIleLysGlu 512
          ||||||:||||||   ::|||||||:||||||||||||:||| |||||:::||||||||||
Qy  1873 GGCTCCATGGAGGGTGCT-----AAGGGTGGCACCGTGCTCCCATCCTAGGGAAGGGC 1926

Db   513 GlyLysLeuAlaValValAlaLeuAspAspAspGlyIleThrGlnPheValValLeuSerAspGly 532
          :||||||:||||::|||||:|||||:|||::|||||:||||||:||:||||||  :||:|||
Qy  1927 GACCTGANCGGC-------TGTCAGCTGCCCCGCGGCGGAGCAGGTGGTAAAGCC 1974

Db   533 SerTyrPheGlyGluIleSerIleLeuAsnIleLysGlySerLysAlaGlyAsnArgArg 552
          :||  :::||:|||||||||||::||||||||||||||:|||:||||:||||:|||||
Qy  1975 AANGCCGACGTGAAGGGCGTGAGGTACTACGTCTCCAGCGTGCTGGCCGCTG 2034

Db   553 ThrAlaAsnIleLysSerIleGlyTyrSerAspLeuPheCysLeuSerLysAspAspLeu 572
          :::|||::|||||:::|||||||:||||||||:||||| |||   |||||||||||||||
Qy  2035 CACGACAGCCTTCGCGCGTGTACCCCGAGTTTGCCCGCGGCGCTTCAGTCGTGTGGGCCTCCGAGGG 2094

Db   573 MetGluAlaLeuThrGlyTyrProAsp----------AlaLysThrMetLeuGlu 587
          :::|||||:|||:||  ::||||:|||       :||::||:|||:||||||:|||
Qy  2095 GAGCTCAGCTACTACAACCTGGGTGCTGCTGCGGGAGGCTTGCGGACCAGCTCCCTG 2154

Db   588 GluLysGlyArgGlnIleIleLeuMetLysAspGlyLeuLeuAspGluAsnIleAlaAsnLeu 607
          |||:::||:|||:||:||||||:::|||||||:::||||:|||::||||||:::|||::|:|||
Qy  2155 ACGGGCGACAATACCCTTATGTCCACGCTGGAGGAGAAGAGACA 2199

Db   608 GlySerAsp----------ProLysAspLeuGluGlyLysValThr 619
          :|||||||       ||||||||:|:||||||||||||||:|||||
```

RESULT 5

```
ID   CNG3_HUMAN     STANDARD;      PRT;   694 AA.
AC   Q16281; Q9UP64;
DT   01-NOV-1997 (Rel. 35, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Cyclic-nucleotide-gated cation channel alpha 3 (CNG channel alpha 3)
DE   (CNG-3) (Cyclic nucleotide gated channel alpha 3) (Cone
DE   photoreceptor cGMP-gated channel alpha subunit).
GN   CNGA3 OR CNG3.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=9517019; PubMed=9517456;
RA   Wissinger B., Muller F., Weyand I., Schuffenhauer S., Thanos S.,
RA   Kaupp U.B., Zrenner E.;
RT   "Cloning, chromosomal localization and functional expression of the
RT   gene encoding the alpha-subunit of the cGMP-gated channel in human
RT   cone photoreceptors.";
RL   Eur. J. Neurosci. 9:2512-2521(1997).
RN   [2]
RP   SEQUENCE OF 320-580 FROM N.A.
RX   MEDLINE=9517019; PubMed=752814;
RA   Distler M., Biel M., Flockerzi V., Hofmann F.;
RT   "Expression of cyclic nucleotide-gated cation channels in non-sensory
RT   tissues and cells.";
RL   Neuropharmacology 33:1275-1282(1994).
RN   [3]
RP   VARIANTS RMCH, AND VARIANT M-153.
RX   MEDLINE=98324775; PubMed=9662398;
RA   Kohl S., Marx T., Giddings I., Jaegle H., Jacobson S.G.,
RA   Apfelstedt-Sylla E., Zrenner E., Sharpe L.T., Wissinger B.;
RT   "Total colourblindness is caused by mutations in the gene encoding the
RT   alpha-subunit of the cone photoreceptor cGMP-gated cation channel.";
RL   Nat. Genet. 19:257-259(1998).
CC   -!- FUNCTION: VISUAL SIGNAL TRANSDUCTION IS MEDIATED BY A G-PROTEIN
CC       COUPLED CASCADE USING CGMP AS SECOND MESSENGER. THIS PROTEIN CAN
CC       BE ACTIVATED BY CYCLIC GMP WHICH LEADS TO A OPENING OF THE CATION
CC       CHANNEL AND THEREBY CAUSING A DEPOLARIZATION OF CONE
CC       PHOTORECEPTORS.
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC   -!- TISSUE SPECIFICITY: PROMINENTLY EXPRESSED IN RETINA.
CC   -!- DISEASE: DEFECTS IN CNGA3 ARE A CAUSE OF ROD MONOCHROMACY (RMCH OR-
CC       RM); ALSO KNOWN AS TOTAL COLORBLINDNESS OR ACHROMATOPSIA. RMCH IS
CC       AN AUTOSOMAL RECESSIVELY INHERITED CONDITION.
CC   -!- SIMILARITY: BELONGS TO THE CYCLIC NUCLEOTIDE-GATED CATION CHANNEL
CC       FAMILY.
CC   -!- DATABASE: NAME=Mutations of the CNGA3 gene;
CC       NOTE=Retina International's Scientific Newsletter;
CC       WWW="http://www.retina-international.com/sci-news/cnga3mut.htm".
```

```
DR   EMBL; AF065314; AAC17440.1; -.
DR   EMBL; S76069; AAD14208.1; -.
DR   Genew; HGNC:2150; CNGA3.
DR   MIM; 600053; -.
DR   MIM; 216900; -.
DR   InterPro; IPR001622; K+channel_pore.
DR   InterPro; IPR000636; M+channel_nlg.
DR   InterPro; IPR000595; cNMP_binding.
DR   Pfam; PF00027; cNMP_binding; 1.
DR   Pfam; PF00520; ion_trans; 1.
DR   SMART; SM00100; CNMP; 1.
DR   PROSITE; PS00888; CNMP_BINDING_1; 1.
DR   PROSITE; PS00889; CNMP_BINDING_2; 1.
DR   PROSITE; PS50042; CNMP_BINDING_3; 1.
KW   Ionic channel; Ion transport; CAMP-binding; Transmembrane;
KW   Multigene family; Vision; Disease mutation; Polymorphism.
FT   TRANSMEM    171    192       POTENTIAL.
FT   TRANSMEM    305    325       POTENTIAL.
FT   TRANSMEM    378    397       POTENTIAL.
FT   NP_BIND     482    605       CAMP.
FT   BINDING     549    549       CAMP (POTENTIAL).
FT   BINDING     564    564       CAMP (POTENTIAL).
FT   VARIANT     153    153       T -> M.
FT   VARIANT     163    163       P -> L (IN RMCH).
FT                                /FTId=VAR_010902.
FT   VARIANT     283    283       R -> Q (IN RMCH).
FT                                /FTId=VAR_010903.
FT   VARIANT     283    283       R -> W (IN RMCH).
FT                                /FTId=VAR_010904.
FT   VARIANT     291    291       T -> R (IN RMCH).
FT                                /FTId=VAR_010905.
FT   VARIANT     410    410       R -> W (IN RMCH).
FT                                /FTId=VAR_010906.
FT   VARIANT     529    529       V -> M (IN RMCH).
FT                                /FTId=VAR_010907.
FT   VARIANT     547    547       F -> L (IN RMCH).
FT                                /FTId=VAR_010908.
FT   VARIANT     557    557       G -> R (IN RMCH).
FT                                /FTId=VAR_010909.
SQ   SEQUENCE   694 AA;  78838 MW;  AE00B4EE760D70A0 CRC64;
```

Alignment Scores:

Pred. No.:		Length:	694
Score:	4.84e-10	Matches:	142
	424.50	Conservative:	95
Percent Identity:	42.25%	Mismatches:	217
Percent Similarity:	25.31%	Indels:	107
Best Local Similarity:	6.97%		
Query Match:			

DB:

US-09-965-830-1_COPY_6_3257 (1-3252) x CNG3_HUMAN (1-694)

```
                    1                    Gaps:    17

Qy    787 GTCGTGACCTGGCCGTGCCGTGAGGTCCTTCACATTGTGCTGAATTCCGTACC 846
          |||  :::  ::::::|||::||:||:||||||::::
Db    205 ValLeuAspTyrSerAlaAspValLeuTyrValLeuAspValLeuValArgValArgThr 224

Qy    847 ACATTCGTGTGTCCAAGTCGGGCCCAGGTGGTGTTTGCCCAAAGTCCACTAC 906
          |||  ::::   :::  |||
Db    225 GlyPheLeu---GluGlnGlyLeuMetValSerAspThrAsnArgLeuTrpGlnHisTyr 243

Qy    907 ---GTCACCACCTGGTTCCTGCTGGATGTCATCGACAGGCTGCCCTTTGACCTGCTACAT 963
          :::  ||||||||:::
Db    244 LysThrThrThrGlnPheLeuLeuAspValIleAspArgLeuProPheAspLeuLeuHisTyr 263

Qy    964 GCCTTCAAGGTCAACGTGTACTTGGGGGCCCATCTGCTGAAGAGGTGGCCGTGGCC 1023
          |||||||  :::    :::|||:::|||
Db    264 ---LeuLysValGlyThrAsnTyrProGlu------ValArgPheAsnArgLeuLeuLys 280

Qy   1024 CTGCTGCGCCTGCTGTTCCGGGCTGTCCGGGACCGGTACTCG 1059
          |||||:::
Db    281 PheSerArgLeuPheGluPhePheAspArgThrGluThrArgThrAsnTyrProAsnMet 300

Qy   1060 --CAGTACAGCGCCGTGGTGGTCGACACTGCCATGGCCGTGTTCGCCCTGCTCGCGCAC 1116
          :::|||||||
Db    301 PheArgIleGlyAsnLeuValLeuTyrIleLeu------IleIleIleHis 315

Qy   1117 TGGGTCGCGTTGCCTGTCTGTTTTACATTGCCAGGGGAGATCGAGCAGCAGAATCCGAG 1176
          |||||  :::::|||:::|||
Db    316 TrpAsnAlaCysIleTyrPheAlaIleSerLys 326

Qy   1177 CTGCCTGACATTGGCTGGCTGCAGGAGCGTGCGCCGCCGACTGGAGACTCCTACACCTG 1236
          :::
Db    327 ------PheIleGlyPhe 330

Qy   1237 GTGGGGCCGGACGGACCAGGAGCTGGACAGCAGCTGCGAAGTGACAAGCTGCAGCAGC 1296
          |||||||
Db    331 ------GlyThrAspSerTrpValTyrProAsnIleSerIlePro 343

Qy   1297 AGGGAGGGCCAACGGGACGGGGCTGGAGAGCTGCGCGGCGGCCTAC 1356
          |||||||
Db    344 GluHisGlyArg------LeuSerArgLysTyr 352

Qy   1357 ATCACCTCCTCTACTTCGCACTCAGCACGCCTCACCAGGTGGGCTTCGGCAAGCTGTCC 1416
          ||||||||||||:::
Db    353 IleTyrSerLeuTyrTrpSerThrLeuThrLeuThrThrIleGly---GluThrProPro 371

Qy   1417 GCCAACACGGACACCGAGAAGATCTTCTCCATCGNCACCATGCTCATCGCGCCCTGANG 1476
          |||||  :::|||
Db    372 ProValLysAspGluGlyTyrLeuPheGluPheTrpAlaAsnLeuLysThrVal 391

Qy   1477 CACGCGGTGGTGTTGGGGAACGTGACGGCCATCANTCCAGGCGATGACCGCCGGCTTT 1536
          |||  ||||||
Db    392 PheAlaThrIleValGlyAsnValGlySerMetIleSerAsnMetAsnAlaSerArgAla 411

Qy   1537 CTGTACACAGCGCCACGCGCGACCTGGGCGACTACACCGCATCCACCACGTACCCCAAG 1596
          :::
Db    412 GlyPheGlnAlaLysIleAspSerIleLysSerGlnTyrMetGlnPheArgLysValThrLys 431

Qy   1597 CCCCTCAAGCACCGGCCATGCTGGGCGGGTGAACAATGGCATC 1656
          |||||:::
Db    432 AspLeuGluThrArgValIleAsnArgTrpPheAspTyrLeuTrpAlaAsnLysLysThrVal 451

Qy   1657 GACACCACGGAGCTGCTGCCAGAGCCTCCGCCAGACATCGCCATGGAC 1716
          ||||||
Db    452 AspGluLysGlyValLeuSerLeuProAspLysLeuLysAlaGluThrIleAlaAsn 471

Qy   1717 CTGTACCAAGGAGGTCCTG---CAGCTGCCACTGTTTGAGGCGGGCCCAGCCGCTGCCTG 1773
          :::
Db    472 ValHisLeuAspThrLeuLysLysValArgIlePheGluGlnCysGluAlaGlyLeuLeu 491

Qy   1774 CGGGCACTGTCTCTGCCCTGCGGCCCGCCCCTTCGCACGCCGGACTACCTCATCCAC 1833
          |||||||
```

```
Db    492 ValGluLeuValIleLeuLysLeuArgProThrValPheSerProGlyAspTyrIleCysLys 511
          |||  ||||||||  :::|||:::||||||:::
Qy   1834 CAAGGCCGATGCCCTGCGAGCCCTCACTTTGTCCTGGCTCCATGGAGGTCCTC--- 1890

Db    512 LysGlyAspIleGlyLysGluMetTyrIleIleLeuAsnGluGlyLysLeuAlaValValAla 531
          |||||::::::|||
Qy   1891 --AAGGGTGGACCGTGCCTCGCCATCTTCAGGAAAGGCGGACCTGACCGC--- 1938

Db    532 AspAspGlyValThrGlnPheValValLeuSerAspGlySerTyrPheGlyGluIleSer 551
          :::|||
Qy   1939 ---TGTGAGCTGCCCCGCGGCAGGTGGTAAAGGCCAATCCGACGTGAAGGGCTG 1995

Db    552 IleLeuAsnIleLysGlySerGlyAsnArgArgThrAlaAsnIleArgSerIle 571
          :::||||||||
Qy   1996 ACGTACTGCGTGCTCCTGCAGTGTCTGCACGAGACCTGCCTTGCGGTTGTGTAC 2055

Db    572 GlyTyrSerAspLeuPheCysLeuSerLysAspAspLeuMetGluAlaLeuThrGluTyr 591
          |||  |||||||:::|||
Qy   2056 CCCGAGTTTGCCCCGCGCCTTCAGTCGTGCGTCGGCGGGAGCTCAGCTACACACCTGGGT 2115

Db    592 ProGlu------- 593
          :::
Qy   2116 GCTGGGGGGCGAGGCTCTGCGAGAGGTGGACACCAGCTCCCTGAGCGGGCGACAAATACCCTTATG 2175

Db    594 AlaLysLysAlaLeuGluLeuGluGlyLysArgGlnIleLeuLeuMetLysAspAspAsnLeuAspLeu 613
          |||||  |||:::
Qy   2176 TCCACGCTGGAGGAGGAGAAGGAGAGACAGATGGGGGACCAGGGAGCCCCAGCCCCA 2235

Db    614 GluGluLeuAlaArgGlnAlaGlyAlaAsp---------ProLysAsp 625
          |||
Qy   2236 GCTGATGAGGCCCTCCAGCCCCGCTGTCCCCCTGGCGTCGCACCTCCTCAATCCTCAGCCTGCC 2295

Db    626 LeuGluGlyLysValGluGlnIleLeuGlySerSerLeuAspThrLeuGlnThrArgPheAla 645
          :::|||||  ||||||
Qy   2296 AAGCTGCTATCCCCACGCTCGAACACAGCACCC--------CGGCCTGCTGTCTCTA------2337

Db    646 ArgLeuLeuAlaGluIleTyrAsnAlaThrGlnMetLysGlnArgLeuSerGlnLeu 665
          :::||||||||::
Qy   2338 ------GGTGGCAGAGGAGGCCAGAGGCCAGGGGGCTTTGAAGGCTGAG 2382

Db    666 GluSerGlnValLysGlyGlyLysProLeuAlaAspGlyGluValProGlyAsp 685
          |||||::|||
Qy   2383 GGT 2385

Db    686 Ala 686
          |||
Db    686 Ala 686
```

RESULT 6

CNG1_CHICK

```
ID   CNG1_CHICK   STANDARD;   PRT;   735 AA.
AC   Q90805;
DT   01-NOV-1997 (Rel. 35, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   01-NOV-1997 (Rel. 35, Last annotation update)
DE   Cyclic nucleotide gated channel, cone photoreceptor, alpha subunit
DE   (CNG channel 1) (CNG-1).
OS   Gallus gallus (Chicken).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC   Gallus.
OX   NCBI_TaxID=9031;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=93264082; PubMed=7684234;
RA   Bonigk W., Altenhofen W., Mueller F., Dose A., Illing M.,
RA   Molday R.S., Kaupp U.B.;
RT   "Rod and cone photoreceptor cells express distinct genes for
RT   cGMP-gated channels.";
RL   Neuron 10:865-877(1993).
CC   -!- FUNCTION: VISUAL SIGNAL TRANSDUCTION IS MEDIATED BY A G-PROTEIN
CC       COUPLED CASCADE USING CGMP AS SECOND MESSENGER. THIS PROTEIN CAN
CC       BE ACTIVATED BY CYCLIC GMP WHICH LEADS TO A OPENING OF THE CATION
CC       CHANNEL AND THEREBY CAUSING A DEPOLARIZATION OF CONE
```

```
CC       PHOTORECEPTORS.
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC   -!- SIMILARITY: BELONGS TO THE CYCLIC NUCLEOTIDE-GATED CATION CHANNEL
CC       FAMILY.
CC       -----------------------------------------------------------------------
CC       This SWISS-PROT entry is copyright. It is produced through a collaboration
CC       between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC       the European Bioinformatics Institute. There are no restrictions on its
CC       use by non-profit institutions as long as its content is in no way
CC       modified and this statement is not removed. Usage by and for commercial
CC       entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC       or send an email to license@isb-sib.ch).
CC       -----------------------------------------------------------------------
DR   EMBL; X89598; CAA61757.1; -.
DR   InterPro; IPR000636; M-channel_nlg.
DR   InterPro; IPR000595; cNMP_binding.
DR   Pfam; PF00027; cNMP_binding; 1.
DR   Pfam; PF00520; Ion_trans; 1.
DR   SMART; SM00100; cNMP; 1.
PROSITE; PS00888; CNMP_BINDING_1; 1.
PROSITE; PS00889; CNMP_BINDING_2; 1.
PROSITE; PS50042; CNMP_BINDING_3; 1.
KW   Ionic channel; Ion transport; cAMP-binding; Transmembrane; Vision;
KW   Multigene family.
FT   DOMAIN        1    210    CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    211    230    H1 (POTENTIAL).
FT   DOMAIN      231    243    EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    244    262    H2 (POTENTIAL).
FT   DOMAIN      263    286    CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    287    306    H3 (POTENTIAL).
FT   DOMAIN      307    344    EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    345    367    H4 (POTENTIAL).
FT   DOMAIN      368    419    CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    420    439    H5 (POTENTIAL).
FT   DOMAIN      440    523    EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    524    544    H6 (POTENTIAL).
FT   DOMAIN      545    735    CYTOPLASMIC (POTENTIAL).
FT   NP_BIND     532    654    CAMP (BY SIMILARITY).
FT   BINDING     591    591    CAMP (POTENTIAL).
FT   BINDING     606    606    CAMP (POTENTIAL).
FT   CARBOHYD    449    449    N-LINKED (GlcNAc...) (POTENTIAL).
SQ   SEQUENCE    735 AA;    85031 MW;    A67ADFD942CEFCE CRC64;
```

Alignment Scores:
```
Pred. No.:
Score:                   5.27e-10
Score:                   423.50
Percent Similarity:      40.74%
Best Local Similarity:   40.74%
Percent Similarity:      22.99%
Query Match:             6.96%
DB:
```

US-09-965-830-1_COPY_6_3257 (1-3252) x CNG1_CHICK (1-735)

```
Qy       363 GGAGGTGGCTGCTCTTCCTAGTCTCTCCACAAGGA                                                                                                 395
             ||| |||:||| ||:|||:::|| ||||||:
Db        67 GlyArgGlyAlaMetAlaArgLeuSerLeuSerPheValValSerLeuArgSerTrpAlaThr 86

Qy       396 ------------CATCACCGAAACCAAGAAACCG------------------AGGGGGCCC 425
                         |||     :::|||    |||    |||||||
Db        87 ArgHisLeuHisHisGlnAspGlnArgProAspSerPheLeuGluArgIleArgGlyPro 106

Qy       426 CGACAGATGGAA------------------GGAGACAGGTGG 449
             ||:||                          ||||||||||
Db       107 GluLeuValGluValSerSerArgGlnSerAsnIleArgSerPheLeuGlyIleArgGlu 126

Qy       450 TGGCCGGCGGCCGGATA-------TGGCCGGCGACGAGATCCAAAGGCTTCAATGCCAAC-- 498
             |||      :::       |||      ::: |||||||| |||
Db       127 GlnProGlyValAsnGlyProTrpPro------LeuAlaArg-PheAsnValAsnPh 144

Qy       499 ------CGGCGGCGGAGCCGGCCGGTGCTCTACCACCTGTC 533
                   :::                               |||
Db       144 eSerAsnAsnThrAsnGluAspLysGluLysGluGluLysLy 164
```

```
Qy       534 CGGGCACCTGCAGAAGCAGCCCAAGGCCAGCACACAAGCTCAATAAGGGGGTGTTTGGGA 593
             :::::::: :: ||| :::::::: ||| :::::: :::
Db       164 sGluGluLysLysGluGluLysGluLysGluGluLysLysGluGluLysAspAspLysLysGluAspLysLys 184

Qy       654 GCACTGTGGGCCACTGAGAGCCCACCTGGGATGGCTTCATCCTGCCACACTCTATGT 713
             :: :::
Db       204 pProSerSerAsnMetTyrTyrTyrAsnTrp------------------LeuThrIleIl 217

Qy       714 GGCTGGTCACTGTGCCCTTACACGTGTGTGTGAGCACAGCACGGGAGCCCAGTGCGCCCG 773
             ||| ||| |||:: ::: |||::: |||
Db       217 eAlaAlaProValPheTyrAsnTrpCysMetLeuIleCysArg--------------------- 231

Qy       774 CGGGCCCCCCCAGCGTCTGT----------------------------------------- 792
             |||
Db       231 ----------------------AlaCysPheAspGluLeuGlnIleAspHisIleLysTrpLeuPh 247

Qy       793 ----------GACCTGGCCGTGGAGGTGCCTCTTCATCCTTGACATTGTGCTGAATTTCCGTACCAC 848
             ||| ||||||||||||:::: ||||||||
Db       247 eLeuAspThrValGlnPheLeuAspAspLeuThrLeuProThrAspLeuAlaTyr-- 267

Qy       849 ATTCGGTGTCAAGTCGGGCCAGGTGGTGTTGTTTGCCCAAAGTCCATTGCCTCCACTACGT 908
             ||||:|| ::: :::||| ||| |||||||||| ||||||||
Db       267 yPheLeu---GluGlnLeuLeuAlaAspGluLysLysLeuArgAspHisIleTyrTh 286

Qy       909 C---ACCACCTGGTTCCTGCTGGATGTCATCCGCAGCCTGCCCTTTGACCTGCTACATGC 965
             ||| ||||||||||||:::::::||||||||||||:::
Db       286 rGlnThrValGlnPheLeuAspValLeuSerLeuLeuProThrAspLeuAlaTyr-- 305

Qy       966 CTTCAAGGTCAACGTGTTACTTCGGGGCCCCATCTGCTGAAGAGAGGGTGGCCCTGCTGCGGCCT 1025
             ::: |||||||||||||
Db       306 -LeuLysLeuGlyLeuAsnTyrProGlu-----LeuArgPheAsnArgGlnLeuLeuArgIl 323

Qy       1026 GCTGCGCCCTGCGTTCCCCGCGGCGTAGCCCGGTACTCG 1059
             : |||||| ::: :::
Db       323 eAlaArgGlnLeuGluPhePheAspAspArgThrArgThrAsnTyrProAsnMetPh 343

Qy       1060 -CAGTACAGCCGCCGCGGCTGGTGGTGCGTGGACACTGCTCATGCCGTTGCCCTGCTGCGCGACTG 1118
             |||||||||
Db       343 eArgIleGlyAsnLeuValLeuTyrIleLeu---------------------IleIleIleHisTr 358

Qy       1119 GGTCGCCTGCGTCTGGTTTACATTGGCCAGCGGGAGGATCGAGAGCAGCACGGAGATCC------- 1173.
             || |||||||::::|||||||||| :::|||
Db       358 pAsnAlaCysIleTyrPheAlaIleSerLysValIleGlyPheGlyThrAspSerTrpva 378

Qy       1174 --------GAGCTGCCTGAGATTGGGCTGGGCTGCAGGAGCCTGGCCCCCGACTGGA 1220
             :::||||||||| ::::
Db       378 lTyrProAsnValSerIleProGluTyrGly--------ArgLeuSerArgLys---- 393

Qy       1221 GACTCCCTACTACCTGGTGGGCGGGCAGCCCAGCTGGAGGGAACAGCTGCCCGGCCAGAGTGA 1280
             |||||||||||||:::::::|||||||
Db       393 -------------------------------------------IleIleIleHisTr 393

Qy       1281 CAACTGCAGCAGCAGCAGCGAGGGCAACGACGAGGGCCTGTGGGGCGGCCCGTC 1340
Db       393 ------------------------------------------------------- 393

Qy       1341 GCTGCGCCAGGCGCCTACATCACCTCCCTCTACTTCGCACTCAGCAGGCCCTCACCAGGTGGG 1400
             :::|||||||||||
Db       394 ------------------TyrIleTyrSerLeuTyrPheArgLeuThrIleLeuThrHisIleLeuGl 409

Qy       1401 CTTCGGCGAACGTGTCCCCACAACAGCAGACACCAGAGAAGACTTCTCAACTATCGTGACCAGGTGGGGGCT 1460
             :::
Db       409 y---GluThrProProProValLysAspGluTyrLeuPheValValIleAspPhePhe 428

Qy       1461 CATCGGCGCCCTCGATGCAGCACGCGGGTCGTGTTTGGAACGTGACGCCCATCATCACGGCCAT 1520
             |:::: |||
Db       428 uValGlyValLeuIlePheAlaThrIleValGlyAsnValGlySerMetIleSerAsnMe 448

Db       1521 GTACGCCCGCGCCGCTTTCTCGTTACCACAGCAGCGCACAGCGGCGGACCTGCGCCCACTCAATCCGCAT 1580
```

Qy 448 tAsnAlaSerArgAlaGlyPheGlnAlaLysValAspSerIleLysGlnThrMetHisPh 468
            |||:::: |||:::
Db 448 CCACGGTATCCCAAGCCCCTCAAGCACGCCATGCTGGAGTACTTCCAGGCCACCTGGGC 1640

Qy 468 eArgLysValThrLysAspLeuGluLuAlaArgValIleLeuTrpPheAspTyrLeuTrpTh 488
            |||::||:| |||:||
Db 468 eArgLysValThrLysAspLeuGluLuAlaArgValIleLeuTrpPheAspTyrLeuTrpTh 488

Qy 1641 GGTGAACAATGGCATCGACACCACCGAGCTGCTGCGAGCCTCCCTGACGAGCTGCGCGGC 1700
            |||| |||::||||
Db 488 rAsnLysThrGluLeuLeuArgAlaSerLeuThrSerCysAlaAl 508

Qy 1701 AGACATCGCCATGCCACCTGCACAAGGAGGTGCCTG---CAGCTGCCACTGTTTGAGGCGGC 1757
            ||:||||||:|::|||||| :||||||||::||
Db 508 aGlyIleIleAlaAsnValHisLeuAspThrLeuLysValLysSerValArgIlePheAspAsp 528

Qy 1758 CAGCCGCGGCGGCTGCCCTGCGCGCACTGTCTCTCTGGCCCCTGCACGGCCGG 1817
            || |||||||||||:|| ||:||||
Db 528 sGluAlaGlyIleLeuLeuIleGluLeuValLeuLysLysLeuProThrValPheSerProGl 548

Qy 1818 CGAGTACCTCATCCACCAAGGGGATGCCCTGCGAGCCCTCTACTTTGTCTGCTGGGTC 1877
            |:||:|||||| ::||||||::||
Db 548 yAspTyrIleCysLysLysGlyAspIleGlyArgGluMetTyrIleIleLysLysGlyLy 568

Qy 1878 CATGGAGGTGCTC-----AAGGGTGGCACCGTGCTCTGCCATCCTAGGGAAGGGCGACCT 1931
            ||||||||||| |||:|||
Db 568 sLeuAlaValAlaAspAspGlyIleThrGlnPheValValLeuSerAspGlySerTyr 588

Qy 1932 GATCGGC-----------TGTGAGCTGGTGCCCCGGCGGAGGAGTGGTAAAGGCCAATGC 1979
            ||||||
Db 588 rPheGlyGluIleSerIleLeuAsnIleLysGlySerLysSerGlyAsnArgArgThrA 608

Qy 1980 CGACGTGAAGGGGCTGACGTACTGCGTTCCTGCAGTGTCTGCACTGGCTGCCTGCACGA 2039
            ||||||||||||::||||
Db 608 aAsnIleArgSerIleGlyTyrSerAspLeuPheCysLeuSerLysAspAspAspLeuMetGl 628

Qy 2040 CAGCCTTGCGCGTGACCCGAG 2061
            |||||||||||||||||
Db 628 uAlaLeuThrGluTyrProGlu 635

RESULT 7

ID   CNG3_MOUSE   STANDARD;   PRT;   631 AA.
AC   Q9JJZ8; Q9WV01;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Cyclic-nucleotide-gated cation channel alpha 3 (CNG channel alpha 3)
DE   (CNG-3) (CNG3) (Cyclic nucleotide gated channel alpha 3) (Cone
DE   photoreceptor cGMP-gated channel alpha-subunit to all cone photoreceptors in the mouse
DE   retina.";
GN   CNGA3 OR CNG3.
OS   Mus musculus (Mouse)
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6; TISSUE=Retina;
RX   MEDLINE=20273944; PubMed=10813773;
RA   Hirano A.A., Hack I., Waessle H., Duvoisin R.M.;
RT   "Cloning and immunocytochemical localization of a cyclic nucleotide-
RT   gated channel alpha-subunit to all cone photoreceptors in the mouse
RT   retina.";
RL   J. Comp. Neurol. 421:80-94(2000).
RN   [2]
RP   SEQUENCE OF 95-631 FROM N.A.
RC   STRAIN=129/Sv;
RX   MEDLINE=9307448; PubMed=10377453;
RA   Biel M., Seeliger M., Pfeifer A., Kohler K., Gerstner A., Ludwig A.,
RA   Jaissle G., Fauser S., Zrenner E., Hofmann F.;
RT   "Selective loss of cone function in mice lacking the cyclic
RT   nucleotide-gated channel CNG3.";
RL   Proc. Natl. Acad. Sci. U.S.A. 96:7553-7557(1999).
CC   -!- FUNCTION: VISUAL SIGNAL TRANSDUCTION IS MEDIATED BY A G-PROTEIN
CC   COUPLED CASCADE USING CGMP AS SECOND MESSENGER. THIS PROTEIN CAN
CC   BE ACTIVATED BY CYCLIC GMP WHICH LEADS TO A OPENING OF THE CATION
CC   CHANNEL AND THEREBY CAUSING A DEPOLARIZATION OF CONE
CC   PHOTORECEPTORS.
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC   -!- TISSUE SPECIFICITY: PROMINENTLY EXPRESSED IN RETINA.
CC   -!- SIMILARITY: BELONGS TO THE CYCLIC NUCLEOTIDE-GATED CATION CHANNEL
CC   FAMILY.

DR   EMBL; AJ243933; CAB89685.1; -.
DR   EMBL; AJ238239; CAB42841.1; -.
DR   EMBL; AJ238240; CAB42891.1; JOINED.
DR   EMBL; AJ238241; CAB42891.1; JOINED.
DR   MGD; MGI:1341818; Cnga3.
DR   InterPro; IPR001622; K+channel_pore.
DR   InterPro; IPR000636; M+channel_nlg.
DR   InterPro; IPR000595; cNMP_binding.
DR   Pfam; PF00027; cNMP_binding; 1.
DR   Pfam; PF00520; ion_trans; 1.
DR   SMART; SM00100; cNMP; 1.
DR   PROSITE; PS00888; cNMP_BINDING_1; 1.
DR   PROSITE; PS00889; cNMP_BINDING_2; 1.
DR   PROSITE; PS50042; cNMP_BINDING_3; 1.
KW   Ionic channel; Ion transport; cAMP-binding; Transmembrane;
KW   Multigene family; Vision.
FT   TRANSMEM   112   133   POTENTIAL.
FT   TRANSMEM   244   264   POTENTIAL.
FT   TRANSMEM   320   340   POTENTIAL.
FT   NP_BIND   423   546   CAMP.
FT   BINDING   490   490   CAMP (POTENTIAL).
FT   BINDING   505   505   CAMP (POTENTIAL).
FT   BINDING   505   505   CAMP (POTENTIAL).
FT   CONFLICT   110   110   L -> Y (IN REF. 2).
FT   CONFLICT   157   157   L -> V (IN REF. 2).
SQ   SEQUENCE   631 AA;   72641 MW; 4FA8CD3B9AA3FE6C CRC64;

Alignment Scores:
Pred. No.:
Score: 5.83e-10
Score: 422.50
Percent Similarity: 43.36%
Best Local Similarity: 25.25%
Query Match: 6.94%
DB:

Length:		631
Matches:		152
Conservative:		109
Mismatches:		224
Indels:		117
Gaps:		22

US-09-965-830-1_COPY_6_3257 (1-3252) x CNG3_MOUSE (1-631)

Qy 334 GATGTGATTACCCATTAAAGAANATGAGAAAAGGGAGGAGGTGGCTCTCCTTCCTAGTCTCTCACAAG 393
            |||:||| :: |||| |||:||||| |||:|||::||||
Db   24 AspLeuAspHisValGluAsnGlyLeuGlyArgValSerArgLeuIleIleSerIleArg 43

Qy 394 --------------------GACATCAGCGGAAACCAAGAACCGAGGGGGCCCGAC 429
                                ||| |||:|||
Db   44 AlaTrpAlaSerArgHisLeuHisAspGluAspGlnThr------------ProAsp 58

Qy 430 AGATGGAAGGAGGACAGGTGGTGGCCGGGGCGGGGCCGATATGGCCGGCGGCGACGATCCAAAGGCTTC 489
            ||||||||||||||| ::||| ||:|||
Db   59 SerPheLeuAsp-------------------ArgPheHisGly-------SerGlnLeuLys 70

Qy 490 AATGCCAACCGCGCGGCGGCGGAGCCGCGGGCGCCGTGCTCTACCACCGTGTCCGGGCCACCTGCAGAAG 549
            ::: ||| ||| ||| |||
Db   71 GluValSerThrArgGluSerAsnAla--------------------------------------- 79

Qy 550 CAGCCCAAG---GGCAAGCACAAGCTCAATAAGGAGGGGTGTTGTTGGGGAGAAACCAAACTTG 606
            ||||||| ||||| ||| ||| ||| |||
Db   80 GlnProAsnProGlyGluGlyLysProProAspGlyGlyGlnLysArgLysGluGlu--- 98

```
Qy   607  CCTGAGTACAAAAGTAGCCGCCATCCGGAAGTCGCCCTTCATCCTGTTGCACTGTGGGGCA  666
                |||                    ::|::||||
Db    99  ------------ProIleValAlaAspProSerSerAsnIleTyrCysArgTrp        112

Qy   667  CTGAGAGCCAACTGGGATGGCTTCATCCTGCTCCCCACTCTATGTGCGTCACTGTG       726
                |||           ||||
Db   113  ----------------------------IleAlaLeuProVal                  120

Qy   727  CCCTACAGCCGGTGTGTGTGACCACAGCACGGGAG---------CCCAGTGCCGGC     771
                |||::::                         |||
Db   121  PheTyrArgAsnTrpCysLeuLeuValCysArgAlaCysPheAspGluLeuGlnSerGlnHis  140

Qy   772  CGGCGCCGCCACCAGGTCTGACCTGGGAGGTCCTCTTCCATCCTTGACATTGTG        831
                ::::|||||||||:::||||:::|||||||:::||||||||||:::::
Db   141  LeuThrLeuThrLeuLeuValLeuAspAlaAspAlaValLeuThrTyrValLeuAspMetLeu 160

Qy   832  CTGAATTTCCGTACCACATTCGTTGTCATACGTGGGGGCCAGTGGGGTGGTGTTGCCCAAAGTCC  891
                ||||||          :::       |||:::   |||
Db   161  ValAlaArgAlaArgThrGlyPheLeu---GluGlnGlyLeuMetValArgAspThrLysArg    179

Qy   892  ATTTGCCTCCACTACGTC----ACCACCTGGTTCCTCCTGGAGTGCAAGACGGCGTGCCC  948
                |||::::                ||||||||||:::|||::::
Db   180  LeuThrProLysHisTyrPheThrLeuHisPheLysAspIleLeuSerLeuIlePro       199

Qy   949  TTTGACCTGCTACATGCCTTCAAGGTACAACGTGTTACTTCGGGGGCCATCCTGCTGAAGACG  1008
                |||||||              :::       |||:::
Db   200  ThrAspAspAlaAlaTyr-------LeuLysLeuGlyValAsnTyrProGlu-----LeuArgPhe 216

Qy  1009  GTGCGGCCTGCTGCGCGACTCGCCCTGCTCGCCTTCCGGCGGCTGGACGGG            1059
                ||||||||||||:::                  :::
Db   217  AsnAsnArgLeuLeuLysPheSerArgLeuPheGluPheAsnArgThrArgThr          236

Qy  1060  CAGTACAGCGGCGGCTGGTGCTGACCACTGCTCATCGGACCGCTGTTCGCCCTG---CTCCGGCCAC  1116
                |||||||             :::       |||
Db   237  AsnTyrProAsnValPheArgIleGlyAsnLeuValLeuTyrThrLeuIleIleHis        256

Qy  1117  TGGGTCGCCTCGCTTCTCGGTTTTTTACATTGGGCAGCGGGAGATGAGCAAGTCCAGGAATCCGAG  1176
                |||:::                 ||||||:::||:::||||:::
Db   257  TrpLysAsnAlaCysIleTyrTyrPheAlaIleSerLys                      267

Qy  1177  CTGCCTGAGAATTGGCCTGACGCTGCAGGAGCCGCGGACACTCCGACTACCTG         1236
                ||||||                        |||
Db   268  -------PheIleGlyPhe                                         271

Qy  1237  GTGCGGCCCGAGGCCAGCTGGAGGGAACAGCTCCGGCCAGAGTGACAACTGCCAGCAGCAG  1296
                |||||||||             :::       |||
Db   272  ----------GlyThrAspSerTrpProValTyrProAsnThrSerLysPro            284

Qy  1297  AGCGGAGGCCAACGGACACGGCGGCTGGCACTGCTCGGGCGCGTCGCGCCACGGTCTAC    1356
                |||                  :::                 |||
Db   285  -------------GlyThrArgAlaArg                                293

Qy  1357  ATGCACCTGCCTCCCTCTACTTGGCCACTCAGCAGCCTCACCACGGTGGGCTTCGGGCAACGTGTCC  1416
                |||:::|||||||                   |||
Db   294  IleTyrTyrSerLeuThrTyrTrpSerThrHisTyr---GluThrProPro              312

Qy  1417  GCCAACACGGACACCGAGAAGATCTTCTCCACATCGCCACCATCGCGGCGCCCTGATG     1476
                ||||||||||||||        :::        :::
Db   313  ProValLysAspGluThrLeuPheLeuIleAlaSerProPheAlaGlnTyrIleLeuIle     332

Qy  1477  CACCGCGGTGGTGTTGTGGGAACGTGACCGCCATCATCCAGCCACATCTAGCGCCCCGGCTTT  1536
                |||                   |||:::
Db   333  PheAlaThrIleLeuGlyAsnValGlySerMetIleSerAsnMetAsnAlaProAspArgVal  352

Qy  1537  CGTGTACCAGCGCCGCCACCCGCGCCATCCANCCACCGGTATCCCCAAG             1596
                |||:::                           |||:::
Db   353  GluPheGlnLnAlaLysIleAsnAspSerValLysGlnThrTyrMetGlnInPheArgLysValThrLys  372

Qy  1597  CCCCTCAAGCCAGCCATGTCTGGAGTACTTCCAGGCCGGGGTGAACAATGGCATC       1656
                |||:::|||||                |||:::
Db   373  AspLeuGluThrArgValIleArgThrTrpPheAspPheTyrLeuTrpAlaAsnArgLysThrVal  392

Qy  1657  GACACCACGGAGCGTGCTGAGAGCCTCCCTGACGAGCACGCCAGACAACGCCANGCAC     1716
```

```
Db   393  AspGlyLysGlyValLeuLeuLysAsnLeuProAspLysLeuLysAlaGlyGluIleAllaIleAsn  412
               |||:::||||:::|||||||||||:::||||:::||||||:::::||||:::||:::
Qy  1717  CTGCACAAGGAGGTCCTG---CAGCTGCCCACTGTTGAGGGCGGCCAGCCGCGGGCCCTG    1773
               |||:::|||||||||:::|||:::||||:::|||||||:::
Db   413  ValHisLysAspThrLeuLysLysValArgIlePheGluGlnLeuProPrySerGluLeuLeu    432

Qy  1774  CGGGCACTGTCTCCTGGGCCCCTGGGGCCCGGCCCTGCACGGCGGGCGAGTACCTCATCCAC   1833
               |||:::||||||||||||:::||||:::|||||||:::||||:::|||:::||||
Db   433  ValGluLeuValLeuLysLeuArgProThrValPheSerProGlyAspTyrIleCysLys       452

Qy  1834  CAAGGCCGATGCCCCTGCAGGCCCCTACTTTGTTGTCCTGGGTCCCATGGAGGTTGCTC---  1890
               |||:::|||||||||:::||||:::|||||||:::||::::||||
Db   453  LysGlyAspIleGlyArgGluMetTyrIleIleLysGluGlyLysLeuAlaValValAla     472

Qy  1891  ---AAGGGTGGCACCGGTGCCACGGTCGCTCGGCCATCCNAGGAGAGGGGGACCTGATCGGC  1938
               :::||||:::|||||||:::||||::::||||||:::||||
Db   473  AspAspGlyValThrGlnPheValValLeuSerAspGlySerTyrPheGlyGluIleSer    492

Qy  1939  ---TGTGAGCTGCCCCGCCGGGGAGCAGGTGGTAAAGGCCAATGCCGAGGTGAAGGGGTG    1995
               :::||||:::||||||:::||||||:::||||||:::
Db   493  IleLeuAsnIleLysGlySerLysSerGlyAsnArgArgThrAlaAsnIleArgSerIle    512

Qy  1996  ACGTACTGCGGTCCTGCCAGTGTCTGCCAGTGGCTCGGCCGGACCGCCTTGCCGTCTGTAC   2055
               |||:::||||:::||||||:::||||:::||||||:::||||:::|||:::
Db   513  GlyTyrSerAspLeuPheCysLeuSerLysSerAspAspAspLeuMetGluAlaLeuThrGluTyr  532

Qy  2056  CCCGAG    2061
               |||:::
Db   533  ProAsp    534
```

RESULT 8

```
CNG_DROME

ID   CNG_DROME               STANDARD;      PRT;     665 AA.
AC   Q24278; Q9V7L5;
DT   01-NOV-1997 (Rel. 35, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Cyclic-nucleotide-gated cation channel (CNG channel).
GN   CNG OR CG7779.
OS   Drosophila melanogaster (Fruit fly).
OC   Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;
OC   Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
OC   Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
OX   NCBI_TaxID=7227;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=95045396; PubMed=7957070;
RA   Baumann A., Frings S., Godde M., Seifert R., Kaupp U.B.;
RT   "Primary structure and functional expression of a Drosophila cyclic
RT   nucleotide-gated channel present in eyes and antennae.";
RL   EMBO J. 13:5040-5050(1994).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Berkeley;
RX   MEDLINE=20196006; PubMed=10731132;
RA   Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA   Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA   George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA   Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA   Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA   Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA   Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA   Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA   Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA   Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA   Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA   Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA   de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA   Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA   Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,
RA   Fosler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA   Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
```

```
RA   Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA   Hostin D., Houston K.A., Howland T.J., Wei M.H., Ibegwam C.,
RA   Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA   Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA   Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA   Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA   Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA   Mount S.M., Moy M., Murphy B., Murphy L., Murray D.M., Nelson D.L.,
RA   Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA   Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA   Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA   Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA   Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA   Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA   Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissenbach J.,
RA   Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA   Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA   Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA   Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RL   "The genome sequence of Drosophila melanogaster.";
RT   Science 287:2185-2195(2000).
CC   -!- FUNCTION: APPROXIMATELY 50-FOLD MORE SENSITIVE TO CGMP THAN TO
CC       CAMP. MAY BE INVOLVED IN TRANSDUCTION CASCADES OF BOTH
CC   -!- TISSUE SPECIFICITY: EXPRESSED IN ANTENNAE AND THE VISUAL SYSTEM.
CC       INVERTEBRATE PHOTORECEPTORS AND OLFACTORY SENSILLAE.
CC   -!- SIMILARITY: BELONGS TO THE CYCLIC NUCLEOTIDE-GATED CATION CHANNEL
CC       FAMILY.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; X89601; CAA61760.1; -.
DR   EMBL; AE003807; AAF58033.1; -.
DR   FlyBase; FBgn0014462; Cng.
DR   InterPro; IPR001622; K+channel_pore.
DR   InterPro; IPR000636; M+channel_nlg.
DR   InterPro; IPR000595; cMP_binding.
DR   Pfam; PF00027; cNMP_binding; 1.
DR   Pfam; PF00520; ion_trans; 1.
DR   SMART; SM00100; cNMP; 1.
DR   PROSITE; PS00888; CNMP_BINDING_1; 1.
DR   PROSITE; PS00889; CNMP_BINDING_2; 1.
DR   PROSITE; PS50042; CNMP_BINDING_3; 1.
KW   Ionic channel; Ion transport; cAMP-binding; Transmembrane.
FT   DOMAIN        1    110       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    111    131       H1 (POTENTIAL).
FT   DOMAIN      132    138       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    139    159       H2 (POTENTIAL).
FT   DOMAIN      160    186       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    187    207       H3 (POTENTIAL).
FT   DOMAIN      208    253       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    254    274       H4 (POTENTIAL).
FT   DOMAIN      275    325       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    326    346       H5 (POTENTIAL).
FT   DOMAIN      347    481       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    482    502       H6 (POTENTIAL).
FT   DOMAIN      503    665       CYTOPLASMIC (POTENTIAL).
FT   CAMP        437    559       CAMP (BY SIMILARITY).
FT   NP_BIND     496    496       CAMP (POTENTIAL).
FT   BINDING     511    511       CAMP (POTENTIAL).
FT   CARBOHYD    135    135       N-LINKED (GLCNAC.   .) (POTENTIAL).
FT   CONFLICT     73     73       G -> R (IN REF. 1).
SQ   SEQUENCE   665 AA;  75823 MW;  6EFFC9A7CA243660 CRC64;
```

Alignment Scores:
Pred. No.:
Score: 6.35e-10   Length: 665
Score: 421.50   Matches: 155
Percent Similarity: 43.27%   Conservative: 102

US-09-965-830-1_COPY_6_3257 (1-3252) x CNG_DROME (1-665)

Best Local Similarity: 26.09%
Query Match: 6.92%
Indels: 110
Mismatches: 227
Gaps: 23

```
Qy  1323 GCTGCTGCGGGGCGCCGTCGCCGCAGCGCCGCTACATCACCTCCCTCTTCGCACTCAG 1382
              ::|||||:
Db   291 ------ArgAsnAsnThrLeuGlnArgGlnThrTyrIleTyrSerPheTyrTrpSerThrIle 308

Qy  1383 CAGCCTCACAGCGGTGGGCCTTCGGCAACGTGTCCCAACACGGACAGGAGAAGATCTT 1442
              :::|||||||:::.|||||    :::||
Db   308 uThrLeuThrThrIleGlyGlnThrProThrProGluAsn---AspValGluTyrLeuPh 327

Qy  1443 CTCCAACTGCACCAATGCTCATCGGCCCTGATGCACGGCGTGGTGGTTTGGAACGTGAC 1502
              .|::                 |::
Db   327 eValValAlaAlaAspPheLeuAlaGlyValLeuIlePheAlaThrIleValGlyAsnIleGl 347

Qy  1503 GGCCATCATCCAGGCGCATGTACGCCCGGCGCTTTCTGTACCACAGCGCGCACCGCGGACCT 1562
              ||||:::
Db   347 ySerMetIleSerAsnMetAsnValAlaAlaArgValGluPheGlnAsnAspGlyMetAspGlyVa 367

Qy  1563 GCGCGGACTACATCACCGTATCCCCAAGCACCCTCAAGCAGGCGGCCATGCTGGAGTA 1622
              |||      :::||||||||||:::||||||||||||||:::
Db   367 lLysGlnThrMetAlaPheArgArgGlyHisGlnLeuGluGlnLeuAlaArgValIleArgTr 387

Qy  1623 CTTCCAGGCCACCTGGGGCGGGGTGAACAATGGCATGACGACCACCGCAGCTGTCGAGAGCCT 1682
              :::||||       :::|||||||||     ||||      .:|||      .||||
Db   387 pPheAlaTyrThrTrpSerGlnSerGlyAlaLeuAspGluGlyArgValLeuAlaAlaAle 407

Qy  1683 CCCTGACGGAGCTGCGCGCCGCAGACAATCGGCCACGACGAGGTCCTG---CACCT 1739
              ||||||||||:::|||||||||:::||||     :::      |||  :::
Db   407 uProAspGlyLeuLysLysAlaGlyIleLeuAlaIleHisMetAspThrLeuLysGlnVa 427

Qy  1740 GCCACTGTTTGAGCGGCGGCGGCGGCGGCGCGGGGGCACGTGCTCCTGGGCCCTGCGGCC 1799
              |||      :::|||||||||||||||:::
Db   427 lArgIlePheHisSerAspThrGluProGlyLeuLeuGluLeuValLeuLysLysLeuLysSe 447

Qy  1800 CCCCTTCTGCACGGGCCGGCTGTACCTCCATCCACCAAGGCAGGCCGCCCTCAGCGCCCTA 1859
              :::|||||||:::      |||||:::      .:|||
Db   447 uGlnValPheSerProGlyAspTyrIleCysArgLysGlyAspValGlyLysGluMetTyr 467

Qy  1860 CTTTGTCGCTGTGGCGTGGCACCGTGCTCGCCAT-----AAGGGTGGCACCGTCTCGCCAT 1913
              |||      :::|||||||||||:::
Db   467 rIleValLysArgGlyLysLeuSerValValGlyIleThrValLeuAlaTh 487

Qy  1914 CCTAGGGAAGGGGGAccTGATCGGC-------TGTGAGCGGCTGCCCCGGCCGGGGGAGCA 1961
              |||||||      |||||||||
Db   487 rLeuGlyAlaGlySerValPheGlyLeuValSerValLeuGluIleLeuAlaGlyAsnArgh 507

Qy  1962 GGTGGTAAAGGCCAATGCCGACGTGAAGGGGCTGACGTACTGCGTCCTGCAGTGTCTGCA 2021
              :::|||||||||||
Db   507 rGlyAsnArggArgThrAlaAsnValArgSerLeuGlyTyrSerAspLeuPheCysLeuAl 527

Qy  2022 GCTGGCCTGGCCTCGCACGACAGCCCTTGCCGTGTACCCCGAG 2061
              |:::||||||||||:::
Db   527 aLysArgAspLeuTrpGluThrLeuSerAspTyrProGlu 540
```

RESULT 9

ID   CNG1_CANFA   STANDARD;   PRT;   691 AA.
AC   Q28279;
DT   01-NOV-1997 (Rel. 35, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   cGMP-gated cation channel alpha 1 (CNG channel alpha 1) (CNG-1) (CNG1)
DE   (Cyclic nucleotide gated channel alpha 1.) (Cyclic nucleotide gated
DE   channel, photoreceptor) (Cyclic-nucleotide-gated cation channel 1)
DE   (Rod photoreceptor cGMP-gated channel alpha subunit).
GN   CNGA1 OR CNCG1 OR CNCG.
OS   Canis familiaris (Dog).
OC   Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
OX   NCBI_TaxID=9615;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Beagle X Briard;
RA   Veske A.;

RL   Submitted (AUG-1996) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=97415722; PubMed=9268598;
RA   Zhang Q., Pearce-Kelling S., Acland G.M., Aguirre G.D., Ray K.;
RT   "Canine rod photoreceptor cGMP-gated channel protein alpha-subunit:
RT   studies on the expression of the gene and characterization of the
RT   cDNA.";
RL   Exp. Eye Res. 65:301-309(1997).
-!- FUNCTION: VISUAL SIGNAL TRANSDUCTION IS MEDIATED BY A G-PROTEIN
    COUPLED CASCADE USING CGMP AS SECOND MESSENGER. THIS PROTEIN CAN
    BE ACTIVATED BY CYCLIC GMP WHICH LEADS TO A OPENING OF THE CATION
    CHANNEL AND THEREBY CAUSING A DEPOLARIZATION OF ROD
    PHOTORECEPTORS.
-!- SUBUNIT: HOMOTETRAMER OR HIGHER OLIGOMER.
-!- SUBCELLULAR LOCATION: Integral membrane protein.
-!- SIMILARITY: BELONGS TO THE CYCLIC NUCLEOTIDE-GATED CATION CHANNEL
    FAMILY.

DR   EMBL; X99914; CAA68186.1; -.
DR   EMBL; U83905; AAB61707.1; -.
DR   PROSITE; PS00888; CNMP_BINDING_1; 1.
DR   PROSITE; PS00889; CNMP_BINDING_2; 1.
DR   PROSITE; PS50042; CNMP_BINDING_3; 1.
DR   SMART; SM00100; cNMP; 1.
DR   Pfam; PF00520; ion_trans; 1.
DR   Pfam; PF00027; cNMP_binding; 1.
DR   InterPro; IPR000595; cNMP_binding.
DR   InterPro; IPR000636; M+channel_nlg.
KW   Ionic channel; Ion transport; cAMP-binding; Transmembrane;
KW   Multigene family; Vision.
FT   DOMAIN   1   163   CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM   164   184   H1 (POTENTIAL).
FT   DOMAIN   185   197   EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM   198   216   H2 (POTENTIAL).
FT   DOMAIN   217   240   CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM   241   260   H3 (POTENTIAL).
FT   DOMAIN   261   298   EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM   299   321   H4 (POTENTIAL).
FT   DOMAIN   322   373   CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM   374   393   H5 (POTENTIAL).
FT   DOMAIN   394   477   EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM   478   498   H6 (POTENTIAL).
FT   DOMAIN   499   691   CYTOPLASMIC (POTENTIAL).
FT   NP_BIND   486   608   CGMP (POTENTIAL).
FT   BINDING   545   545   CGMP (POTENTIAL).
FT   BINDING   560   560   CGMP (POTENTIAL).
FT   CARBOHYD   424   424   N-LINKED (GLCNAC) (POTENTIAL).
SQ   SEQUENCE   691 AA; 80251 MW; 0775CAA42F065275 CRC64;

Alignment Scores:
Pred. No.:   6.33e-10   Length: 691
Score: 421.50   Matches: 145
Percent Similarity: 42.54%   Conservative: 126
Best Local Similarity: 22.76%   Mismatches: 265
Query Match: 6.92%   Indels: 101
                                          Gaps: 17

DB:
US-09-965-830-1_COPY_6_3257 (1-3252) x CNG1_CANFA (1-691)

```
Qy  343 CCCATAAAGAATAGAGAAA----------GGGGAGGTGGCTCTCGTACCTTCCTAGTCTCTCACAAG 393
             |||   ::    |||||||||||:::::
Db   74 ProSerGlnArgGluGlnThrTyrLeuProGlyAlaIleAlaLeuPheAsnValAsnAsnSer 93

Qy  394 GACATCACGCGAAACCAGGAAACCGAGGGGGCCCCGACAGATGGAAGGAGGAGCAGGTGGTGGC 453
```

```
Db    94  SerAsnLysGluGlnGluProLysGluLysLysLysLysGluLys----------           110
               :::
Qy   454  CGGGCCCGGATATGCCGGCGGCAAAAGGCTTCAATGCCACCAGCCGGCCCGG              513
               ::::::::::::::::|||::::::::||||||||||||
Db   111  ----------LysSerLysSerGlyAspLysAsnLysLysLysLysAspSerGlu           125

Qy   514  GCCGTGCTCTACCACTGTCCGGGGCGACCTGCAGAAGCAAGCCAAGCACAAGCTC          573
               :::::||:::    |||       |||        |||
Db   126  ------------LysLysLysLysLysLysGlyLysLysLys                        137

Qy   574  AATAAAGGGGGTGTTTGGGAGGGAGAAACAAACTTGCCTGAGTACTACAAAGTAGCCGCCCATCCGG  633
               |||||    |||        |||     |||       :::
Db   138  AsnLysGlyGlyLysGlyLysLysAspLysLysGluGlu----------Lys             151

Qy   634  AAGTCGCCCTTCAATCCCGTTGCACTGTGGGGGCCACTGAGAGCCACCTGGGCTTCATC       693
               |||:::   :::|||   :::|||||||||||:::
Db   152  LysSerValMetValIleAspProAlaGlyAsnMetTyrTyrAsnTrpLeuPheCysIle      171

Qy   694  CTGGTCCGCCACACACTCTATGTGGCGTGTCACTGTCCCTACAGCCGTGTGTGTGAGCACAGCA   753
               |||       |||:::|||            |||        |||:::|||
Db   172  ThrLeuProValMetTyrArgAsnTrpThrMetValIleLeuAlaArgAlaCysPheAspGluLeu  191

Qy   754  CGGGAGCCCAGTGCGCCGCCGCCGCCGGCTGTGAACCTGGCGGCGGAGGTCCTC           813
               :::||||||:::||||||                :::|||||||||:::
Db   192  GlnSerAspTyrLeuGluThrTyrTrp-----IleIlePheAspTyrLeuLeuSerLeuAspIleVal  209

Qy   814  TTCATCCTTGACATTGTGCTGAATTCCGTGTCCAAGTGCGGCCAGGTG                 873
               |||:::   |||||||||:::||||||
Db   210  TyrLeuLeuAspMetPheValArgThrArgThrGlyTyrLeuGluGlnGlyLeuLeuVal      229

Qy   874  GTGTTTGCCCCAAAGTCCATTGCCTCCACTACGTCACCACCTGGTTCCTGCTGGGATGTC      933
               |||       :::|||        :::|||        :::
Db   230  ArgGluGluValLysLysLeuIleGlyLysTyrLysSerAsnLeuGlnPheLysLeuAspPhe    249

Qy   934  ATCCAGCAGCCTGCCCTTTGACCTGCTACAATGCCTTCAAG---GTCAAC---GTGTACTTCGGG  990
               |||:::||||||||||           :::|||||||||
Db   250  LeuSerValIleProThrAspPheLeuLeuTyr---PheLysLeuGlyTrpAsnTyrProGlu    268

Qy   991  GCCCATCTGCTGAAGCACGGTGCGCCGCCCTGCGCCTGCGCTTCCGGCGGCTGGAC         1050
               ::|       |||||||||        |||       |||:::
Db   269  IleArgLeuAsnArgLeuLeuArgIleSerArgMetPheGluPhePheGlnArgThrGlu      288

Qy  1051  CGGTACTCGCAGTACAGCAGCCGTGGGTG---CTGACACTGCTCATGGCCGGTGTTCGCCCTG   1107
               :::|||       |||:::|||          |||:::||||||
Db   289  ThrArgThrArgAsnTyrProAsnIlePheArgIleSerAsnLeuValMetTyrIleValIle    308

Qy  1108  CTCGCGGCGACTGGGTGCGCCGTCGGTGTTTTACATTGGCCAGGGGATCGAGAGCAGC       1167
               |||||||||       |||||||||:::||||||
Db   309  IleLeuHisTrpAsnAlaCysValTyrPheSerIleSerLysAlaIleGlyPheGlyAsn      328

Qy  1168  GAATCCGAGCTG-------CCTGAGATTGGCTGGCTGGCCGAGGACTGGGC             1209
               |||||||||      :::|||||||||
Db   329  AspThrTrpValTyrProAspValAlaAsnAspProGluPheGly-------ArgLeuAla     345

Qy  1210  CGGCGACTGGAGACTCCCTACTACCTGGTGGGCCGGAGGCCAGCTGGAGGAACAGCTCC      1269
               :::||||||:::||||||
Db   346  ArgLys-------TyrValTyrSerLeuTyrTrpSerThrLeuThrLeu                 359

Qy  1270  GGCCAGAGTGACAACTGCAGCAGCAGGAGCGGCAACGGACGGGGCTGGAGCTGCTG         1329
               |||
Db   347  ------------                                                      347

Qy  1330  GCGGGCCCGTGCCTGCGCGAGCGGCCTACACATCACCTCCTACTTCGCACTCAGCAGCCTG    1389
               |||
Db   348  ------------                                                      347

Qy  1390  ACCAGCGTGGGGCTTCGGGAACGTGTCCGCACACAGACACCGAGAAGAATCTTCTCCATC     1449
               |||:::||:::||        :::||||||        :::
Db   360  ThrThrIleGly---GluThrProProValArgAspSerGlnThrTyrValPheValVal      378

Qy  1450  TGCACCAATGCTCATCGGGCGCCCCTGATGACGCGGCGGTGGTTTGGGAAGCTGACGGCCATC  1509
               |||||||:::|||       ::     ||||||
```

```
Db   379  ValAspPheLeuIleGlyValLeuIlePheAlaThrIleValAlaGlyAsnIleGlySerMet   398
               |||
Qy  1510  ATGCAGCCGCCATGTACGCCCCGGCGCCCTTTCTTGGTACCACCAGCCGACCGGGGCGAC    1569
               :::||||||:::||||||        |||:::||||||:::
Db   399  IleSerAsnMetAsnAlaAlaArgAlaGluProAspGlnIleAlaArgIleAlaTyrLysGln   418
               |||
Qy  1570  TACATCCGCATCCACCGTATCCCCAAGCCCCTCAAGCACAACGGCAACAATGCTGGAGTACTTCCAG  1629
               |||||||||:::       |||||||||:::||||||:::
Db   419  TyrMetHisPheArgAsnValSerLysLysAspMetGluLysArgValIleIleLysTyrPheAsp   438
               |||
Qy  1630  GCCACCTGGGCGGTGAACAATGGCATCGACACACCCGAGCTGCTGCAGAGCCTGCCCTGAC    1689
               :::||||||:::||||||        |||:::||||||
Db   439  TyrLeuTrpThrAsnLysLysThrValAspGluLysGluLysIle--LeuLeuLysTyrLeuAsp   458
               |||
Qy  1690  GAGCTGCGGCGCAGACATGACGCCATGCACCGAACGCACGAAGGAGGTCCTG---CAGTTGCCACTG  1746
               :::||||||:::||        |||:::||||||:::
Db   459  LysLeuArgAlaGluIleIleAsnGluAlaAlaHisLeuAsnAspThrLeuLysLysValArgIle   478
               |||
Qy  1747  TTTGAGCGGCGGCACCAGCCGCCGCCGGGGTGCGCCGGGCCACTGTTCTGGCCCTCGGCCCGCCCTTC  1806
               :::||||||:::||||||        |||
Db   479  PheAlaAspCysGluAlaGlyLeuLeuValGluLeuValLeuLysLeuGlnProGlnVal        498
               |||
Qy  1807  TGCACGCGCGGGGAGCGGGAGTACCTCATCCACCAAGGGCCATGCCGCTGCGAGGCCCTCTACTTTGTC  1866
               :::||||||||||||:::||||||:::
Db   499  TyrSerProGlyAspTyrIleCysLysLysGlyAspIleGlyArgGluMetTyrIleIle        518
               :::
Qy  1867  TGCTCTGGCTCCATGGAGGTGCTC-----AAGGGTGGCACCGTGCGCTCGGCCATCCTAGGG    1920
               |||||||||:::||||||:::
Db   519  LysGluLeuLysLeuAlaValAlaAspAspGlyTyrSerArgLeuPheCysLeuSerLysAspAsp  538
               :::
Qy  1921  AAGGGCGACCTGATCGGC---------TGTGAGCGTGGCCCGCGGGGGGGGGGGGTA        1968
               |||||||||:::
Db   539  AspGlySerTyrPheGlyGluIleSerLeuLeuAsnIleIleGlySerLysAlaGlyAsn      558
               :::
Qy  1969  AAGGCCAATGCCACGTGCAAGGGGCTGAGCGTACTGCGGCTCGTCCCTGCCAGTGTGTCGCTGGCT  2028
               |||||||||:::||||||
Db   559  ArgArgThrAlaAsnIleLysSerIleGlyTyrSerAspLeuPheCysLeuSerLysAsp       578
               |||
Qy  2029  GGCCTGCACGACGACCGCCTTGCGCGTCGTACCCCGAGTTGCCCGCCGGCCTTCAGTCGTCGGGCCTC  2088
               :::||||||:::||||||        |||
Db   579  AspLeuMetGluAlaLeuThrGluTyrProGlu-----------AlaLysThrMet           593
               :::
Qy  2089  CGAGGGGAGCTCAGCTACAACCTGGGTGGTGCTGGGGGGAGGCCTCGCAGAGGTGGACACCAGC  2148
               :::||||||:::||||||
Db   594  LeuGluGlyLysGlyLysIleLeuLeuIleLeuMetLysLysAspGlyLeuLeuAsnIleAla    613
               |||
Qy  2149  TCCCTGAGCGGCGGCGACAAATACCCTATTGTCCACGCTGGAGGAGGAGGAGACA         2199
               |||||||||:::
Db   614  AsnAlaLeuGlySerAsp--------ProLysAspLeuGluGlyLeuGluLysValThr        627
```

```
RT   rod photoreceptor cGMP-gated cation channel.";
RL   J. Biol. Chem. 267:6257-6262(1992).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Kidney;
RX   MEDLINE=95315239; PubMed=7540868;
RA   Karlson K.H., Ciampolillo-Bates F., McCoy D.E., Kizer N.L.,
RA   Stanton B.A.;
RT   "Cloning of a cGMP-gated cation channel from mouse kidney inner
RT   medullary collecting duct.";
RL   Biochim. Biophys. Acta 1236:197-200(1995).
CC   -!- FUNCTION: VISUAL SIGNAL TRANSDUCTION IS MEDIATED BY A G-PROTEIN
CC       COUPLED CASCADE USING CGMP AS SECOND MESSENGER. THIS PROTEIN CAN
CC       BE ACTIVATED BY CYCLIC GMP WHICH LEADS TO A OPENING OF THE CATION
CC       CHANNEL AND THEREBY CAUSING A DEPOLARIZATION OF ROD
CC       PHOTORECEPTORS.
CC   -!- SUBUNIT: HOMOTETRAMER OR HIGHER OLIGOMER.
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC   -!- TISSUE SPECIFICITY: ROD CELLS IN THE RETINA AND INNER MEDULLA OF
CC       KIDNEY.
CC   -!- SIMILARITY: BELONGS TO THE CYCLIC NUCLEOTIDE-GATED CATION CHANNEL
CC       FAMILY.
CC   -------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   -------------------------------------------------------------------------
DR   EMBL; M84742; AAA37425.1; -.
DR   EMBL; U19717; AAA85702.1; -.
DR   EMBL; U19715; AAA85700.1; -.
DR   EMBL; U19716; AAA85701.1; -.
DR   MGD; MGI:88436; Cncg.
DR   InterPro; IPR000836; M-channel_nlg.
DR   InterPro; IPR000595; cNMP_binding.
DR   Pfam; PF00027; cNMP_binding; 1.
DR   Pfam; PF00520; Ion_trans; 1.
DR   SMART; SM00100; cNMP; 1.
DR   PROSITE; PS00888; CNMP_BINDING_1; 1.
DR   PROSITE; PS00889; CNMP_BINDING_2; 1.
DR   PROSITE; PS50042; CNMP_BINDING_3; 1.
KW   Ionic channel; Ion transport; cAMP-binding; Transmembrane;
KW   Multigene family; Vision.
FT   DOMAIN        1    156       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    157    177       H1 (POTENTIAL).
FT   DOMAIN      178    190       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    191    209       H2 (POTENTIAL).
FT   DOMAIN      210    233       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    234    253       H3 (POTENTIAL).
FT   DOMAIN      254    291       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    292    314       H4 (POTENTIAL).
FT   DOMAIN      315    366       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    367    386       H5 (POTENTIAL).
FT   DOMAIN      387    470       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    471    491       H6 (POTENTIAL).
FT   DOMAIN      492    684       CYTOPLASMIC (POTENTIAL).
FT   NP_BIND     479    601       CGMP (POTENTIAL).
FT   BINDING     538    538       CGMP (POTENTIAL).
FT   BINDING     553    553       CGMP (POTENTIAL).
FT   CARBOHYD    417    417       N-LINKED (GLCNAC...) (POTENTIAL).
FT   CONFLICT    113    114       NK -> I (IN REF. 1).
FT   CONFLICT    200    200       D -> N (IN REF. 1).
FT   CONFLICT    576    576       A -> V (IN REF. 1).
FT   CONFLICT    635    635       R -> C (IN REF. 1).
SQ   SEQUENCE   684 AA;  79460 MW;  13EC2A405B6B8CBF CRC64;
```

```
Alignment Scores:
Pred. No.;
Score:                  7.91e-10
Percent Similarity:     419.00
                        42.01%

Length:                 684
Matches:                153
Conservative:           139
```

```
Best Local Similarity:  22.01%
Query Match:            6.88%            Mismatches:   277
DB:          1                           Indels:       126
                                         Gaps:         22

us-09-965-830-1_copy_6_3257.(1-3252) x CNG1_MOUSE (1-684)
```

```
Qy    181 ATGGCAGGCGGGGCTGTGGCCTGCTCCTTCCTTATGGCGCCCAGGACCACCAGTGAGCTCCGGC   240
              :::::   |||              |||        |||   |||       |||
Db     30 MetGluAsnGly--AlaCysSerSerPheSerAspProAsnGlySerLeuSerGlu         48

Qy    241 CAACAGATCCGACAAGGGCCCTGGAGCACGACAGCACAAAGGTGTGAGCTGTACTCTGTAC   300
                |||            |||      |||        |||         |||
Db     49 Glu------------SerGluAsnGluAspSerPhePheArgSerAsn-------SerTyr  62

Qy    301 CGGGAAGAGCGGGGCTCCCGTTCGGGTGTCCCGTGGATGGTGATACCCATAAAGAAGAGAA   360
              ::::::  |||          |||
Db     63 LysArgArgGly------------                                     72

Qy    361 -----GGGGAGGTGGCTCTCTTCCTTAGTCTCTCACAAGGACATCAGCGAAACCAAG      411
               |||||     :::|||||||          |||         |||
Db     73 HisLeuProGlyThrMetALaLeuPheAsnValAsnAsn------SerSerAsnLys    89

Qy    412 AACCGGAGGGGGCCCCCGACAGATGGGAGGAGGAGACAGGTGGTGGGCCGCGGGCCGCCGG   471
              ::::||||:::             |||         |||
Db     90 AspGlnGluProLysGlyLysLysLys                                 104

Qy    472 GCACGATCCAAAAGGCTTCAATGCCAACGCGCGGCGGGAGCCGGGCCGTGCTCTACCACTG   531
              |||||||:::          |||          |||
Db    105 LysLysGluGluLysThrLysGluLysLysGluGluLysGluValValValIle      118

Qy    532 TCCGGGGCCACCTGCAGAAGCAGCCCAAGGAGCAGCAGAGCTCAATAAAGGGGGGTGTTTGGG 591
              ::::||||:::                    |||
Db    119 -----------LysLysLysLysGluGluLysLysLys                      130

Qy    592 GAGAAACCAAACTTGCCTGAGTACAAAGTAGCGCGCCGAAGTCGCCCTTCATCCTG      651
              :::::||||:::                         |||
Db    131 LysLysGluGluLysThrLysGluLysLysGluGluLysGluLysLys           150

Qy    652 TTGCACTGTGTGGGGCACTGAGAGCCCACCTGGGAAGCGGCTTCATCCTGCTCGCCACACTCTAT 711
              :::::::||||:::
Db    151 AspProSerGlyAsnThrTyrTyrTyrAsnTrpLeuProLeuThrLeuProValMetTyr     170

Qy    712 GTGGCTGTCACGTGCCCTACAGCGGTGTGTGTGAGCACAGGAGCCCAGTGCGCC        771
              ||| :::                         |||
Db    171 AsnTrpThrMetIleIleAlaArgAlaCysPheAspGluLeuGlnSerAspTyrLeuGlu    190

Qy    772 CGCGGCGCCGCCCAGCGTCTGTGACCTGGGGAGGTCCTCTTCATCCTTGACATTGTG     831
              :::::::::::::::::::::             |||
Db    191 TyrTrp-----LeuIlePheAspTyrValSerAspValTyrLeuAlaAspMetPhe    208

Qy    832 CTGAATTCCGTACCACATTCGTCCAGTCGGACCGGCAGGTGGGTGTTGCCCAAAGTCC    891
              ::::::||||:::
Db    209 ValArgThrArgThrGlyTyrLeuGluGlnGlyLeuLeuValLysAspArgMetLysLeu  228

Qy    892 ATTTGCCTCCACTACGTTCACCACCTGGTTCTCTGCTGGGTCATCGACCGCCCTCCCTTT  951
              ||| :::                              |||
Db    229 IleGluLysTyrLysAlaAsnLeuGlnPheLysLeuAspValLeuSerValIleProThr  248

Qy    952 GACCTGCTACATGCCTTCAAGGTCAACGTGTACTTCGGG------GCCCAT         996
              ||||||||||||:::                      |||
Db    249 AspLeuLeuTyr------------IleLysPheGlyTrpAsnTyrProGluIleArg     263

Qy    997 CTGCTGAAGACGGTGCCCCTGCCCCTGCTGCGCGCCTGGACCGGTAC            1056
              ||| :::                               |||
Db    264 LeuAsnArgLeuArgIleSerArgMetPheGlnArgPheGluThrArg             283

Qy   1057 TCGCAGTACAGCGCCGCGGTGGTG---CTGACACTGCTCATGGCCGGTCGGGTTCGCCCCTCGCGG 1113
              :::|||                                               |||
Db    284 ThrAsnTyrProAsnIlePheArgIleSerAsnLeuValMetTyrIleValIleIleIle       303

Qy   1114 CACTCGGGTCGCCTGGGTTGTCGTCTGGTTTTACATTGCCACGGGAGATCGAGAGGAGGAATCC  1173
              |||||||:::                                            |||
Db    304 HisThrProAsnAlaCysValTyrTyrSerIleSerGlyLysAlaIleGlyPheGlyAsnAspThr 323
```

```
Qy   1174 GAGCTG------------------------------CCTGAGATTGGCTGCTTGCTGCGAGGAGGCTGGCCCGCCCGA 1215
             |||                                    :::
Db    324 TrpValTyrProAspValAlaAsnAspProGluPheGly------ArgLeuAlaArgLys 340

Qy   1216 CTGGAGACTCCCTACTACCTGGTGGCCCGGAGGGCCAGCTCCGGCCAG 1275
Db    340 ------ 340

Qy   1276 AGTGACAACTGCAGCACGAGCACCAAGCGAGCCCAAGGGAGGAGGCGGC 1335
Db    340 ------ 340

Qy   1336 CCGTGCCTGCCTGCCCAGCCGCCTACATCACCTCCCTCTACTTCGCACTCAGCAGC 1395
Db    340 ------ 340

Qy   1396 GTGGGGCTTCGGCAACGTGTCCGCCACACGAGAAGATCTTCTCCATCTGCACC 1455
             |||:::         ||||||||:::        |||:::
Db    341 ------TyrValTyrSerLeuTyrTrpSerThrLeuThrThr 354

Qy   1456 ATGCTCAATCGGCCCCTCGATGCACGCGCGGTGGTGTTTGGGAACCGCCATCATCCAG 1515
             :::    ||||||||||||:::       |||:::
Db    355 IleGly---GluThrProProProValLeuAspSerGluThrTyrIlePheValValAlaAsp 373

Qy   1516 CGCATGTACGCCCGCGGCTTTCTGTACCACGCGCGGCGACCCGGCGGCGACTCACATC 1575
             |||   ||||||   |||||||:::||||||
Db    374 PheLeuIleGlyValLeuIlePheAlaThrIleValAlaGlyAsnIleGlySerMetIleSer 393

Qy   1576 CGCATCCACCGTATCCCCAAGCCCCTCAAGCACCGCCATCGCCGGAGTACTCAGGCCACC 1635
             |||:::||||:::     ||||||||||:::
Db    394 AsnMetAsnAlaAlaArgAlaGluPheGlnSerArgValAspAlaIleLysGlnTyrMet 413

Qy   1636 TGGGGCGGTGAACAAATGGCATCGAACACCAAGCCGAGCTGCTGCAGAGCCCCTCACGGACTG 1695
             |||:::  |||||||||:::|||||||||||:::|||:::
Db    414 AsnPheArgAsnValSerLysAspMetGluLysArgValIleLysTrpPheAspTyrLeu 433

Qy   1696 CGCGGAGACAACGCCACTGCGACCAGGGCACCTGTTTGAG 1752
             |||||||:::||||||||||:::|||:::
Db    434 TrpThrAsnLysLysThrValAspGluValAlaArgTyrLeuProGlnValTyrSerLeu 453

Qy   1753 GCGGCCAGCCGCGGCCTCGCGGGGACTGTCTCTGGCCCGTGCGGCCCGGCCTTCTCACG 1812
             |||||||   ||||||   |||||||:::||||||
Db    454 ArgAlaGluIleIleAlaAsnValHisLeuAspThrLeuLysLysValArgIlePheAla 473

Qy   1813 CCGGGCGGAGTACCTCATCCACCAAGCGGATGCCCCTGCAGGCCCCTCACTTTGTCTGCTCT 1872
             |||||||||::||||::|||:::   ::::||||
Db    474 AspCysGluAlaGlyLeuLeuValGluLeuValLeuLysLeuGlnProGlnValTyrSer 493

Qy   1873 GGCTCCACTGGAGGTGCTC------AAGGGTGGCACCGTGCTCGCCATCCTAGGGAGGGC 1926
             |||||||||::::|||||:::        :::|||||||   :::|||:::
Db    494 ProGlyAspTyrTyrIleCysLysLysGlyLysAspIleGlyArgGluMetTyrIleIleLysGlu 513

Qy   1927 GACCTTGGATCGGC------TGTGACTGCCCGCGGGACCAGGTGGTAAAGGCC 1974
             |||                :::
Db    514 GlyLysLeuAlaValValAlaLeuAlaAspSerGlyTyrIleThrGlnAsnProValLeuSerAspGly 533

Qy   1975 AATGGCGACGTGAAGGGGCTGACGTACTGCGTGTCTGCCAGTGTCTGGCCTG 2034
             :::|||||||:::   |||:::|||||||||||:::
Db    534 SerTyrPheGlyGluIleSerIleLeuAsnIleLysGlySerLysAlaGlyAsnArgArg 553

Qy   2035 CACGACAGCCTTGGGCGTGTACCCCGCGCTTCAGTGCTGCCTCCGAGGG 2094
             |||:::|||||||   ||||||:::|||:::
Db    554 ThrAlaAsnIleLysSerLeuGlyTyrSerAspLeuPheCysLeuSerLysAspAspLeu 573

Qy   2095 GAGCTCAGCTACTACAACCTGGGTGCTGCTGCAGAGGTGGACAGCAGCCAGCTCCCTG 2154
             |||:::|||||||||:::||||||
Db    574 MetGluAlaLeuThrGluTyrProAsp------AlaLysThrMetLeuGlu 588

Qy   2155 AGCGGCGACAATACCCTPATGTCCAGCTGGTGGAGAGGAGGACA 2199
             |||:::|||||||||:::|||:::
Db    589 GluLysGlyArgGlnIleLeuMetLysAspGlyLeuLeuAspIleAsnIleAlaAsnMet 608

Db    609 GlySerAsp------ProLysAspLeuGluGluLysValThr 620
```

RESULT 11
CNG_ICTPU

ID   CNG_ICTPU   STANDARD;   PRT;   682 AA.
AC   P55934;
DT   01-NOV-1997 (Rel. 35, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   01-NOV-1997 (Rel. 35, Last annotation update)
DE   Cyclic-nucleotide-gated cation channel.
OS   Ictalurus punctatus (Channel catfish).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Siluriformes;
OC   Ictaluridae; Ictalurus.
OX   NCBI_TaxID=7998;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Olfactory neuroepithelium;
RX   MEDLINE=92110008; PubMed=1370374;
RA   Goulding E.H., Ngai J., Kramer R.H., Colicos S., Axel R.,
RA   Siegelbaum S.A., Chess A.;
RT   "Molecular cloning and single-channel properties of the cyclic
RT   nucleotide-gated channel from catfish olfactory neurons.";
RL   Neuron 8:45-58(1992).
CC   -!- FUNCTION: THIS CYCLIC NUCLEOTIDE-GATED CHANNEL IS ACTIVATED
CC   EQUALLY WELL BY BOTH CAMP AND CGMP.
CC   -!- TISSUE SPECIFICITY: OLFACTORY NEURONS.
CC   --
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --
DR   EMBL; M83111; -; NOT_ANNOTATED_CDS.
DR   InterPro; IPR000636; M+channel_nlg.
DR   InterPro; IPR001201; PAP_25A_core.
DR   InterPro; IPR000595; cNMP_binding.
DR   Pfam; PF00027; cNMP_binding; 1.
DR   Pfam; PF00520; ion_trans; 1.
DR   SMART; SM00100; cNMP; 1.
DR   PROSITE; PS00888; CNMP_BINDING_1; 1.
DR   PROSITE; PS00889; CNMP_BINDING_2; 1.
DR   PROSITE; PS50042; CNMP_BINDING_3; 1.
KW   Ionic channel; Ion transport; cAMP-binding; Transmembrane;
KW   Olfaction.
FT   DOMAIN   1   136   CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM   137   157   H1 (POTENTIAL).
FT   DOMAIN   158   169   EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM   170   190   H2 (POTENTIAL).
FT   DOMAIN   191   218   CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM   219   239   H3 (POTENTIAL).
FT   DOMAIN   240   272   EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM   273   293   H4 (POTENTIAL).
FT   DOMAIN   294   311   CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM   312   332   H5 (POTENTIAL).
FT   DOMAIN   333   343   EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM   344   364   H6 (POTENTIAL).
FT   DOMAIN   365   682   CYTOPLASMIC (POTENTIAL).
FT   NP_BIND   455   577   CAMP (BY SIMILARITY).
FT   BINDING   514   514   CAMP (POTENTIAL).
FT   BINDING   529   529   CAMP (POTENTIAL).
SQ   SEQUENCE   682 AA; 78020 MW; 2C78597DC2C74F75 CRC64;

Alignment Scores:
Pred. No.:
Score: 1.18e-09   Length: 682
Percent Similarity: 414.50   Matches: 170
Best Local Similarity: 39.55%   Mismatches: 127
Query Match: 22.64%   Conservative: 127
                      6.81%   Indels: 177
DB: 1   Gaps: 25

US-09-965-830-1_COPY_6_3257 (1-3252) x CNG_ICTPU (1-682)

```
Qy   238  CGCCAACAGATCCGCAAGGCCCTGGACGAGGACAAGGAGTTCAAGGCTGAGCTGATCCTG  297
               ::|||    |:|||  :::::
Db    15  ArgLeuSerValArgSerArgLeuGluGlyGluAlaGlu---ArgAlaGluSerAlaIle   33
          ArgLeuSerValArgSerArgLeuGluGlyGluAlaGlu

Qy   298  TACCGGAAGACGGGG                                               311
          |||
Db   311  ------------------------------------------------------------  310
                                                                  LysThrLeuSerPheCysTyr

Qy   313  CTGCCCGGTTCTGGGTCGTCGGATGGATACCCATAAAGAATGAGAAAGGGGAGGTGGGT  372
          |||||||
Db    34  SerArgThrArgGlyAspGlyAspSerProGlyGluLeuGlnLeuArgValThrAlaLeuGlu   53

Qy   373  CTCTTCCTAGTCTCT---------                                      
               :::||||||:|||
Db    54  LeuProSerAlaLeuMetLeuGlnAlaPheThrGlnArgArgProLeuAlaArgLeuVal   73
                                                    CACAAGGACATCAGCGAACCAAGAACCA  1476

Qy   418  GGGGGCCCGACAGATGGAAGGGAGAGACAGGTGGCTGGCCCGGCGGCCCGGGCACGA  477
              |||||
Db    74  AsnLeuValLeuSerLeuArgGluThrProAlaHisLysSerSerLeuValGlyThrGluGlnArg   93

Qy   478  TCCAAAGGCTTCAAATGCCAACCCGGCGGCGGGGAGGCGGGGGTCGCTGCTACCACCTGTCCGGG  537
             ::|||
Db    94  -----ProAspSerPheLeuGlu   100
                                    CACAAGGACATCAGCGAACCAAGAACCA

Qy   538  CACCTGCAGAAGCAGCCCAAGGCACCAAGCTCAATAAAGGGGGTGTTTGGGGAGAAA  597
             |||
Db   101  PheArgGlyProGlnAlaAlaAsnAspGlnSerAlaAla   113

Qy   598  CCAAACTTGCCTGAGTACAAAAGTAGCCCCGAAGTGCCCATCCGGAAGTCGCCCTTCATCTGTGTTGCAC  657
             :|:
Db   113  ------------------------------------------------------------  113

Qy   658  TGTGGGGGCCAGGCTGGATGGCTTCATCCGTGCCCACACTCTATGTGGGCT  717
          |||
Db   114  ProAlaAlaProAlaProLysGlyLysThrPheLysLysTrpGlyLyrPheValSerGln   133

Qy   718  GTCACGTGCCCTACAGCGTGTGTGGAGCACAGCACGGAGACGCCGAGTCCGCGCGC---  774
          |||
Db   134  SerAspAspIleTyrTyrTyrTyrTrpLeuPheIleAlaLeuAlaSerLeuTyrAsnTrp   153

Qy   775  ---GGCCCGCCACGCTGTGACCTGGGCGGAGGTCCTCTTCATCGTTGGTGTTG  831
             :::||
Db   154  IleLeuMetLeuValAlaArgGlnAlaCysPheAspGluAsnPhePheLeuTrp   173

Qy   832  CTGAAATTCCGTACCACACATTCGTGTCCAAGTCGTGGTTTGCCCCAAAGTCC  891
          ::::::|||||
Db   174  ValGlyLeuAspTyrTyrLeuCysAsp---   189
                                       ValIleTyrIleLeuAspThrCys

Qy   892  ATTTGCCTCCACTACGTC---ACCACCTTGGTTCTCGGATGTCATCGGAGCGGCTGCCC  948
          :::||||
Db   190  IleLeuArgLeuArgThrGlyTyrLeu---GlyLeuGlnGlyLeuGlnLeuValLysAspLeuAlaLys   208

Qy   949  TTTGACCTGCTACATGCCTTCCAAGGTCACGTGTACTTCGGGCCCCATCTGCTGAGACG  1008
          ::::||||
Db   209  LeuArgAspAsnThrTyrIleArgThrLeuGlnPheLysLeuGluSerIleLeuPro   228

Qy  1009  GTGCGCCCTGCTGCGGCCCTGCTTCCGCGGGCTGGACCGG-----  1059
          :::|||
Db   229  ThrGluLeuLeu---------PhePheValThrGlyTyr---   244
                              ValProGlnLeuArgPhe

Qy  1060  CAGTAC--AGGCGCGTGGTTGCTCGACACTCCTCATGGGCGGTCGTCGCCCTGCTGGCCAC  1116
          |||:::
Db   245  AsnArgLeuLeuLeuArgGluMetThrGluPhePheAspAspArgThrGlnThrArgThr   264

Qy  1117  TGGGTCGCCTGCGGCCTGCGGCTGTTTACATTGGCCAAGATCGAGAGACAGCAGGAATCCGAG  1176
          ||||||||||:::||||||::::||::::::
Db   265  AsnTyrProAsnAlaPheArgIleCysAsnLeuLeuTyrIleLeuValIleIleHis   284

Qy  1177  CTGCCTGAGAATGGCTGGCTGCCCGGACTGGAGACTCCCTACTG  1236
          |||:::
Db   285  TrpAsnAlaCysIleLeuTyrTyrAlaIleSerLeuGlyLeuSerSerAspThr---   303

Db   304  ------------------------------------------------------------  306
          --------TrpValTyr
```

```
Qy  1237  GTGGGCCGGAGGCCGAGCTGGAGGGAACAGCCTCCGGCCCAGAGTGACAACTGCAGCAGC  1296
          |||||
Db   307  ------------------------------------SerGlyGlnAsn   310

Qy  1297  AGCGAGGCCAAACGGGACGCGGGCCTGGAGCGTGCTGGGCGGCGGCGGCCCCTAC  1356
             :::|||
Db   311  ------------------------------LysThrLeuSerPheCysTyr   317

Qy  1357  ATCAACCTCCCTCACTTCGCACTCAGCACCAGCCTCACCAGGTGGCTTCGGCAACGGTGG  1416
             :::|||||:::||:
Db   318  ValTyrCysPhePheTyrTrpSerThrLeuThrLeuThrThrIleLeuGluMetProPro   337

Qy  1417  GCCAACACCGACACCGAGAAGATCTTCTCCACATCTCGCACCCATGCTCATCGGCGCCGATG  1476
             ::::||||:::||
Db   338  Val Lys--AspGluGlnTyrValPheValAlaGlyAsnSerArg   356

Qy  1477  CACGCGGTGGTGTTTGGGGAACGTGAACGGCCATCATCCACCGGCATGTACCCGCCCTTT  1536
          ||||||
Db   357  PheAlaThrGluValGlyAsnValGlySerMetIleAlaSerAlaMetAsnAlaThrArgAla   376

Qy  1537  CTGTACCACACCGACACCGAGAAGATCTTCTCCACATCTCGCACCCATGCTCATCGGCGCCAAG  1596
             :::|||||
Db   377  GluPheGlnThrArgIleAspPheValPheValValPheAspLysLysSerHisTyrMetHisPheArgLysValAlaAsn   396

Qy  1597  CACCGGCGGTGGTGTTTGGGGAACGTGAACGGCCATCATCCACCGGCATGTACCCGCCCATC  1656
          ||:::
Db   397  ThrLeuGluThrArgValAlaIleLysTrpPheHisAspTyrLeuTrpThrAsnLysLysThrVal   416

Qy  1657  GACACCACCGAGCTGCTACGAGAGCCTCCCGAGACGCTGGGGCCGGCGGCCGAGAGCATGCACAC  1716
          |||
Db   417  AspGluGlnIleValLeuLeuLysLysSerAsnLeuProAspSerLeuArgArgAlaGlullIleAlaIleAsn   436

Qy  1717  CTGCACAAGGAGGAGTTCCTG---CAGTGCCAACATGTTTGAGGGGCGGGGCCAGCCGCCGCCTG  1773
          ||||||
Db   437  ValHisIleLeuAspThrLeuLysLysValArgIleLeuProGlnAspCysGlyAlaGlyLeuLeu   456

Qy  1774  CGGGCACTGTCTGTCTCTGGCCCTCGGGCGCGCGCCGCCGCCTTGCGCGCGGCCGAGTACCTCATCCAC  1833
          :::|||
Db   457  ValGlyLeuValLeuLeuLysArgProGlnValThrTyrSerProGlyTyrIleCysArg   476

Qy  1834  CAAGGCGGCATGCCCGTGCAGGCCCCTCTCACTTTTGTCGTCTCGGCTCCCATGGAGGTGCTC---  1890
          |||:::
Db   477  LysGlyLysAspIleGlyLysGluMetTyrIleLeuLysGlyLeuGlyGlnLeuAlaValValAlaAla   496

Qy  1891  ---AAGGGTGGCACCGTGCTCTGCCGCTGCCCATCTTAGGGAAGGGGCGGACCTGATCGGC  1938
          :::|||
Db   497  AspAspAspLysValThrGluLeuPheCysGluLeuArgLeuArgLeuAlaGlyIleGlyCysPheGlyLeuSer   516

Qy  1939  ---TGTGAGCGTCCCCGGGGGAGCAGGTGGGTAAAGGCCAATGCCGACGTGAAGGGGCTG  1995
          :::||||
Db   517  IleLeuAsnIleGlnIleGlySerLysMetGlyAsnArgAspArgThrAlaAlaAsnIleArgSerIle   536

Qy  1996  ACGTACTACGGTCCGTCCGCAGTGTGTTCTGCAGCCTGCCTGGCCGCCACGACAGCCTTGCGGTGTAC  2055
          ::::|||||
Db   537  GlyThrTyrSerAspIleLeuPheCysSerLeuSerLysAspAspAspLeuMetCysLeuAlaValAlaValAlaGluTyr   556

Qy  2056  CCCGAGTTGCCCCGCGGCTTCAGTCGTGGCCTCCGGGGGAGCTCAGTCGACTACAACCTGGGT  2115
             :::||||||
Db   557  ProAsp---AlaGlnLysValLeuGlnLeuLeuValArgGlyLysLeuArgGlyLysGlnGly   575

Qy  2116  -------------------------------GCTGGGGAGGGCTTGCAGGAGGTGGACACC---  2145
             ::::|||||
Db   576  LeuLeuAspGluSerValAlaAlaLaGlyLyrLeuGlyValIleAspThrGluGlyLysVal   595

Qy  2146  ------------------------------------AGCTCCCTGACCGGCGAC  2163
             :::|||
Db   596  GluArgLeuAspAlaSerLeuAspIleLeuGlnThrArgPheAlaArgLeuLeuGlyGlu   615

Qy  2163  -----------------------------------------------------------  2163
Db   616  PheThrSerThrGlnArgArgLeuLysGlnArgLeuIleThrAlaLeuGlnLeuCys   635

Db  2164  AATACCCTTATGTCCACGCCTGGAGGAGGAGGAGAAAGGAGACAGATAGGGAGGAGGCCCACGGTC  2223
```

```
Db      636 HisThrGlyGlyLeuGlyLeuLeuSerAspAsnGluAlaGluGlyGluHisAlaGlyValPro 655
                :::|||  ||    :::|    |||       |||||||
Qy     2224. TCCCAGCCCAGCTGATGAGCCCTCCAGCCCC 2256
                 |||||||     |||||||   |||
Db      656 ThrHisThrHisAlaAspIleHisAlaGlnPro 666
```

RESULT 12

```
ID   CNG1_BOVIN     STANDARD;      PRT;    690 AA.
AC   Q00194;
DT   01-APR-1993 (Rel. 25, Created)
DT   01-APR-1993 (Rel. 25, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   cGMP-gated cation channel alpha 1 (CNG channel alpha 1) (CNG1)
DE   (Cyclic nucleotide gated channel alpha 1 (CNG channel alpha 1) (CNG1)
DE   channel, photoreceptor) (Cyclic nucleotide-gated cation channel 1)
DE   (Rod photoreceptor cGMP-gated channel alpha 1 subunit).
GN   CNGA1 OR CNCG1 OR CNCG.
OS   Bos taurus (Bovine).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
OC   Bovidae; Bovinae; Bos.
OX   NCBI_TaxID=9913;
RN   [1]
RP   SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RC   TISSUE=Retinal rod cell;
RX   MEDLINE=90098076; PubMed=2481236;
RA   Kaupp U.B., Niidome T., Tanabe T., Terada S., Boenigk W.,
RA   Stuehmer W., Cook N.J., Kangawa K., Matsuo H., Hirose T., Miyata T.,
RA   Numa S.;
RT   "Primary structure and functional expression from complementary DNA
RT   of the rod photoreceptor cyclic GMP-gated channel.";
RL   Nature 342:762-766(1989).
RN   [2]
RP   3D-STRUCTURE MODELING OF 485-610.
RX   MEDLINE=95365381; PubMed=7543681;
RA   Henn D.K., Baumann A., Kaupp U.B.;
RT   "Probing the transmembrane topology of cyclic nucleotide-gated ion
RT   channels with a gene fusion approach.";
RL   Proc. Natl. Acad. Sci. U.S.A. 92:7425-7429(1995).
RN   [3]
RP   TOPOLOGY.
RX   MEDLINE=92112723; PubMed=1370452;
RA   Wohlfart P., Haase W., Molday R.S., Cook N.J.;
RT   "Antibodies against synthetic peptides used to determine the topology
RT   and site of glycosylation of the cGMP-gated channel from bovine rod
RT   photoreceptors.";
RL   J. Biol. Chem. 267:644-648(1992).
RN   [4]
RP   TOPOLOGY.
RX   MEDLINE=92256398; PubMed=1316156;
RA   Kumar V.D., Weber I.T.;
RT   "Molecular model of the cyclic GMP-binding domain of the cyclic GMP-
RT   gated ion channel.";
RL   Biochemistry 31:4643-4649(1992).
CC   -!- FUNCTION: VISUAL SIGNAL TRANSDUCTION IS MEDIATED BY A G-PROTEIN
CC       COUPLED CASCADE USING CGMP AS SECOND MESSENGER. THIS PROTEIN CAN
CC       BE ACTIVATED BY CYCLIC GMP WHICH LEADS TO A OPENING OF THE CATION
CC       CHANNEL AND THEREBY CAUSING A DEPOLARIZATION OF ROD
CC       PHOTORECEPTORS.
CC   -!- SUBUNIT: HOMOTETRAMER OR HIGHER OLIGOMER.
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC   -!- TISSUE SPECIFICITY: ROD CELLS IN THE RETINA.
CC   -!- SIMILARITY: BELONGS TO THE CYCLIC NUCLEOTIDE-GATED CATION CHANNEL
CC       FAMILY.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
```
```
DR   EMBL; X51604; CAA35947.1; -.
DR   PIR; S07103; S07103.
DR   InterPro; IPR000636; M+channel_nlg.
DR   InterPro; IPR000595; cNMP_binding.
DR   Pfam; PF00027; cNMP_binding; 1.
DR   Pfam; PF00520; ion_trans; 1.
DR   SMART; SM00100; cNMP; 1.
DR   PROSITE; PS00888; CNMP_BINDING_1; 1.
DR   PROSITE; PS00889; CNMP_BINDING_2; 1.
DR   PROSITE; PS50042; CNMP_BINDING_3; 1.
KW   Ionic channel; Ion transport; cAMP-binding; Transmembrane;
KW   Multigene family; Vision.
FT   DOMAIN        1    162       CYTOPLASMIC (PROBABLE).
FT   TRANSMEM    163    183       H1 (PROBABLE).
FT   DOMAIN      184    196       EXTRACELLULAR (PROBABLE).
FT   TRANSMEM    197    215       H2 (PROBABLE).
FT   DOMAIN      216    239       CYTOPLASMIC (PROBABLE).
FT   TRANSMEM    240    259       H3 (PROBABLE).
FT   DOMAIN      260    297       EXTRACELLULAR (PROBABLE).
FT   TRANSMEM    298    320       H4 (PROBABLE).
FT   DOMAIN      321    372       CYTOPLASMIC (PROBABLE).
FT   TRANSMEM    373    392       H5 (PROBABLE).
FT   DOMAIN      393    476       EXTRACELLULAR (PROBABLE).
FT   TRANSMEM    477    497       H6 (PROBABLE).
FT   DOMAIN      498    690       CYTOPLASMIC (PROBABLE).
FT   NP_BIND     485    607       CGMP (POTENTIAL).
FT   BINDING     544    544       CGMP (POTENTIAL).
FT   BINDING     559    559       CGMP (POTENTIAL).
FT   CARBOHYD    423    423       N-LINKED (GLCNAC...) (POTENTIAL).
SQ   SEQUENCE   690 AA;  79602 MW;  A01CFB6567424455 CRC64;
```

US-09-965-830-1_copy_6_3257 (1-3252) x CNG1_BOVIN (1-690)

```
Alignment Scores:
Pred. No.:
Score:                                  1.29e-09
                                        413.50
Percent Similarity:                     40.188
Best Local Similarity:                  22.378
Query Match:                            6.798
DB:
```

Length: 690
Matches: 147
Conservative: 117
Mismatches: 254
Indels: 139
Gaps: 19

```
Qy     400 AGGGAAAACCAAGAAACCGAGGGGGCGCCCCCAGCAGATGCGAAGGAGGAGACA---------GGTGGTGGC 453
               ||||||||||| ||::|||                                      ||| ||| |||
Db      52 SerGluThrGluAsnProHisAlaAlaArgAspSerPheArgSerAsnThrHisGlySerGly 71

Qy     454 ---------CGGCGCCGATATGGCCGGGGACCGCACGATCCAAAGGCTTCAAATGCCAACCGG 501
                          ||| :::||| |||                              ||||||  |||
Db      72 GlnProSerGlnArgGlnGlnTyrLeuProGlyAlaIleAlaLeuPheAsnValAsnAsn 91

Qy     502 CGGCGGAGCCGGCCGTGCCTCTACCACCTGTTCGGGCAACCTGCGAAGCAGCCCAAGGC 561
               ||| ||||| :::|||  |||                              ||| ||| |||
Db      92 ---------SerSerAsnLysGluGlnProLysGlu 101

Qy     562 AAGCACAAGCTCAATAAGGGGGGTGTTTGGGGAGGAAACCAAAACTTGCCTGAG--------- 612
               |||||   ||| :::||| ||||||    ||||||
Db      72 GlnProSerGlnArgGlnGlnTyrLeuProGlyAlaIleAlaLeuPheAsnValAsnAsn 91

Qy     102 LysLysLysLysLysLysLys------ 117
               |||::||  :::|||          -----GluLysSerLysProAspAspLysAsn 117
Db     118 GluAsnLysLysAspProGlyLysLysLysGlyLysGlyLysAspLysLysLys 137

Qy     613 ----------------TACAAAGTAGCCGCCATCCGGAAGTCGCCCTTCATCCCTGTG 654
                               :::|||
Db     612 ----
Db     138 LysGlyGlyLysGlyLysAspLysLysGlyLysGlyLysGluValValValIleAsp 157

Qy     655 CACTGTGGGCCACTGGAGCCCACCTGGATGGCTTCATCCTGCCCCACACTCTATGTG 714
               ||| :::||| ||| |||  ||| |||  :::|||
Db     158 ProSerGlyAsnThrTyrTyrAsnTrpLeuPheCysIleThrLeuProValMetTyrAsn 177

Qy     715 GCTGTGCACTGTTGCCCTACACGCTGTGTGTGAGCACAGCACGGGAGCCCAGTGCCGCGCC 774
               :::||
```

```
Db    178 TrpThrMetIleIleIleAlaArgAlaCysPheAspGluLeuGlnSerAspThrLeuGluTyr 197
          |||
Qy    775 GGCCCGCCCAGGTGCTCTTGACTCTGGCCGCTGGAGCTCCTTCATCCTTGACAATTGTGCTG 834

Db    198 Trp-----LeuAlaPheAspThrTyrLeuSerAspValValTyrLeuLeuLeuMetPheVal 215
          ::::|||||||:||::::|||||:::
Qy    835 AATTTCCGTACCACATTCGTATGGTCCTCAAGGTGGTGTTGTTTCCCAAAGTGCCATT 894

Db    216 ArgThrArgThrGlyTyrLeu---GluGlnGlyLeuLeuValValLysGluGluLeu 234
          |||:::
Qy    895 TGCCTCCACTACGTCACCACCTGG-----TTCCTGCTGAGTCATCGCACGCGGCCTGCCCTTT 951

Db    235 IleAspLysTyrGlySerThrPheGlnPheLysLeuAspValLeuSerValIleLeuProThr 254
          |||
Qy    952 GACCTGCTACATCGCCTTCAAAGGTCACAGTGGCTT-----GCCCAT 996

Db    255 AspLeuLeuTyr------------IleLysPheGlyTrpAsnTyrProGluLeuArg 269
          |||||||
Qy    997 CTGCTGAAGACGGTGCCCTGCTCCTGGTGCTCCAGGGCACCGGTGGACCGGGTAC 1056

Db    270 LeuAsnArgLeuLeuArgIleSerArgMetPheGluPhePheLysLeuAspValLeuSerValIleLeuPro 289
          ::::
Qy    1057 TCCGCAGTACAGGCGCCGTGGTG----CTGACACTGCTCATGGCCGGTGTTCGGCCCTGCTCGCCG 1113

Db    290 ThrAsnTyrProAsnIlePheArgIleSerArgLeuValMetThrTyrIleIleIleIleIle 309
          :::
Qy    1114 CACTGGGTCGCCTGCCTGCCTCTGGTTTTTACATTGGCACAGGGAGATCGAGAGCGAGCGAAATC 1173

Db    310 HisIleTrpAsnAlaCysValTyrPhe------ 317
          |||:::
Qy    1174 GAGCGTGCCTGAGATTGGCTGGCTGCCCGGCCCGGCGACTGCGAGACTCCCTACTAC 1233

Db    317 ------ 317
          :::
Qy    1234 CTGGTGGGCCCGGAGGCAACAGCTCCGGCGAGTGACAACTGCACCAGC 1293

Db    318 ------Ser 318
          |||
Qy    1294 AGCAGGCGAGGCCAACGGACGGGTCGTGGAGTCGTGCTGGGGCCGTCG 1341

Db    318 ------ 318
          |||||
Qy    1342 -----CTGGGCAGGCGGCCTACATCACCTCCCTCTACTTCCGCCACTCAGCAGCCTC 1389

Db    319 IleSerLysSerAlaIleGlyPheGlyAsnAspThrTrpValTyrProAspValAsnAspPro 338
          |||||||||||
Qy    1390 ACCAGGTGGCGGCTTCGGCAACGTGTCCGCCAACAGGCACGAGAGATCTTTCTCCATC 1449

Db    339 AspPheGlyArgLeuAlaArgLysSerTyrValTyrArgSerLeuTyrTrppSerThrLeuThrLeu 358
          :::|||
Qy    1450 TGCACCAANGCCTCATCGCGCCCCTGATCGACGCGGTGGTGTTGGGACGTGAGGGCCATC 1509

Db    359 ThrThrIleGly---GluThrProProProValArgAspSerGluTyrPheEvalVal 377
          ::::|||
Qy    1510 ATCCAGCCATGCTACGCCCGCCGGCTTCTGTCGTACCACAGCCGCGACGCCGGCGAC 1569

Db    378 AlaAspPheLeuLeuGlyValLeuLeuPheLeuAlaThrIleValGlyAsnIleGlySerMet 397
          |||||||
Qy    1570 TACATCCGCCATCACCGTATCCCCAAGCAGCCGCATGCTGGAGTACTCCAG 1629

Db    398 IleSerSerAsnMetAsnAlaLeuAlaArgAlaGluGlnGlnAlaArgIleAspAlaLeuLysGln 417
          |||:::::
Qy    1630 GCCACCTGGGCGGGTGAACAAATGGCCATCGACGAGCCCTGCAGAGCCTCCTGAC 1689

Db    418 TyrMetHisPheArgAsnValSerLysGlyAspMetGluLysLysArgValIleLysTrpPheAsp 437
          |||||||
Qy    1690 GAGCTGCCGCAGACACATGCCACCTGCACAAGGAGGTCCTG---CAGCTGCCACTG 1746

Db    438 TyrLeuTrpThrAsnLysLysSerValAspSerGluArgGluValLeuLysLysTyrLeuProAsp 457
          |||
Qy    1747 TTTGAGGCGGGCCAGCGCGGCGCGGCCACGGCCCTGCTCGCGGGCCCGGCCCTTC 1806

Db    458 LysLeuLysArgAlaGluLeuAlaIleAlaAsnValHisLeuAspThrLeuSerLeuLysValArgIle 477
          :::|||||||:::
Qy    1747 ... 1806

Db    478 PheAlaLeuAspCysGlyLysGlyAlaGlyLeuLeuLeuValGluLeuValLeuLysSerLeuGlnProGlnVal 497
```

```
Qy   1807 TGCACGCCCGGCGGGACGAGTACCTCATCCACCAAGGCGACGATGCCCTGCAGGCCCTCTAC 1866
          ::::||||||||||||::::                                 :::||||
Db    498 TyrSerProGlyAspTyrIleCysLysLysGlyAspIleGlyArgGluMetTyrIleIle 517

Qy   1867 TGCTCTGGCCTGCATGGAGGTGCTC-----AAGGGTGGCACCGTGCTCGCCCATCCTAGG 1920
          :::||||||||::::                        ||||:::
Db    518 LysGluGlyLysLeuAlaValAlaAlaAspAspGlyIleThrGlnPheValValLeuSer 537

Qy   1921 AAGGGCGACCTGATCGGC------------TGTGAGCTGCCCCGGGGGGAGCAGGTGGTA 1968
          |||                  ::::
Db    538 AspGlySerTyrPheGlyIleGluSerIleLeuAsnIleLysGlySerLysAlaGlyAsn 557

Qy   1969 AAGGCCAATGCCGACGTGAAGGGCGTGACGTACTGCGTCCTCCAGTCGTCCGAGCTGGCT 2028
          :::                ::::|||                        ::::
Db    558 ArgArgThrAlaAsnIleLysGlyTyrSerAspAspLeuProGlyLysLysAsp 577

Qy   2029 GGCCTGCACGAGCAGCCCTTGCGTGTACCCGGGCAGTTTGCCCGCGGGTTCGCCCTC 2088
          :::||||||                        |||:::
Db    578 AspLeuMetGluAlaLeuGluThrGluTyrProAsp------------------AlaLysGlyMet 592

Qy   2089 CGAGGGGAGCTCGAGCTACAACCTGGGGTGCCTGGGGAGGGCGCTCTGCAGAGGTGGACACCAGC 2148
          ::::|||||||:::
Db    593 LeuGluGluLysGlyLysGlnIleIleLeuMetLysGlyLeuLeuAspMetAsnIleAla 612

Qy   2149 TCCCTGAGCGCGGACAATACCCTTATGTCACCAGCGTGGAGGAGAGGACA 2199
          ::::||||||:::
Db    613 AsnAlaGlySerAsp------------ProLysAspLeuGluGluLysValThr 626
```

RESULT 13
CNG1_HUMAN
ID   CNG1_HUMAN   STANDARD;   PRT;   686 AA.
AC   P29973; Q16485; Q16279;
DT   01-APR-1993 (Rel. 25, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   cGMP-gated cation channel alpha 1 (CNG channel alpha 1) (CNG channel alpha 1) (CNG-1)
DE   (Cyclic nucleotide gated channel alpha 1 (CNG channel 1) (Cyclic
DE   nucleotide gated cation channel 1)
DE   channel, photoreceptor) (Cyclic-nucleotide-gated cation channel 1)
DE   (Rod photoreceptor cGMP-gated channel alpha subunit).
GN   CNGA1 OR CNCG1 OR CNCG.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Retina;
RX   MEDLINE=92210603; PubMed=1372902;
RA   Pittler S.J., Lee A.K., Altherr M.R., Howard T.A., Seldin M.F.,
RA   Hurwitz R.L., Wasmuth J.J., Baehr W.;
RT   "Primary structure and chromosomal localization of human and mouse
RT   rod photoreceptor cGMP-gated cation channel.";
RL   J. Biol. Chem. 267:6257-6262(1992).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Retina;
RX   MEDLINE=93256211; PubMed=7532814;
RA   Dhallan R.S., Macke J.P., Eddy R.L., Shows T.B., Reed R.R.,
RA   Yau K.-W., Nathans J.;
RT   "Human rod photoreceptor cGMP-gated channel: amino acid sequence,
RT   gene structure, and functional expression.";
RL   J. Neurosci. 12:3248-3256(1992).
RN   [3]
RP   SEQUENCE OF 313-573 FROM N.A.
RX   MEDLINE=95175019; PubMed=7532814;
RA   Distler M., Biel M., Flockerzi V., Hofmann F.;
RT   "Expression of cyclic nucleotide-gated cation channels in non-sensory
RT   tissues and cells.";
RL   Neuropharmacology 33:1275-1282(1994).
RN   [4]
RP   VARIANT ARRP PHE-316, AND VARIANTS GLN-28 AND ASN-114.
RX   MEDLINE=96036047; PubMed=7479749;

```
RA   Dryja T.P., Finn J.T., Peng Y.-W., McGee T.L., Berson E.L., Yau K.-W.;
RT   "Mutations in the gene encoding the alpha subunit of the rod
RT   cGMP-gated channel in autosomal recessive retinitis pigmentosa.";
RL   Proc. Natl. Acad. Sci. U.S.A. 92:10177-10181(1995).
-!- FUNCTION: VISUAL SIGNAL TRANSDUCTION IS MEDIATED BY A G-PROTEIN
CC       COUPLED CASCADE USING CGMP AS SECOND MESSENGER. THIS PROTEIN CAN
CC       BE ACTIVATED BY CYCLIC GMP WHICH LEADS TO A OPENING OF THE CATION
CC       CHANNEL AND THEREBY CAUSING A DEPOLARIZATION OF ROD
CC       PHOTORECEPTORS.
-!- SUBUNIT: HOMOTETRAMER OR HIGHER OLIGOMER. FORMS HETEROOLIGOMERIC
CC       COMPLEX WITH CNG4.
-!- SUBCELLULAR LOCATION: Integral membrane protein.
-!- TISSUE SPECIFICITY: ROD CELLS IN THE RETINA.
-!- DISEASE: DEFECTS IN CNGA1 ARE A CAUSE OF AUTOSOMAL RECESSIVE
CC       RETINITIS PIGMENTOSA (ARRP). A DISEASE THAT LEADS TO DEGENERATION
CC       OF RETINAL PHOTORECEPTOR CELLS.
-!- SIMILARITY: BELONGS TO THE CYCLIC NUCLEOTIDE-GATED CATION CHANNEL
CC       FAMILY.
-!- DATABASE: NAME=Mutations of the CNGA1 gene;
CC       NOTE=Retina International's Scientific Newsletter;
CC       WWW="http://www.retina-international.com/sci-news/cngalmut.htm".
```

```
DR   EMBL; M84741; AAA52010.1; ALT_INIT.
DR   EMBL; S42457; AAB22778.1; -.
DR   EMBL; S76062; AAD14206.1; -.
DR   PIR; A42161; A42161.
DR   Genew; HGNC:2148; CNGA1.
DR   MIM; 123825; -.
DR   InterPro; IPR000636; M+channel_nlg.
DR   InterPro; IPR000595; cNMP_binding.
DR   Pfam; PF00027; cNMP_binding; 1.
DR   Pfam; PF00520; ion_trans; 1.
DR   SMART; SM00100; cNMP; 1.
DR   PROSITE; PS00888; CNMP_BINDING_1; 1.
DR   PROSITE; PS00889; CNMP_BINDING_2; 1.
DR   PROSITE; PS50042; CNMP_BINDING_3; 1.
KW   Ionic channel; Ion transport; cAMP-binding; Transmembrane;
KW   Multigene family; Vision; Disease mutation; Polymorphism;
KW   Retinitis pigmentosa.
FT   DOMAIN         1    160   CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM     161    181   H1 (POTENTIAL).
FT   DOMAIN       182    194   EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM     195    213   H2 (POTENTIAL).
FT   DOMAIN       214    237   CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM     238    257   H3 (POTENTIAL).
FT   DOMAIN       258    295   EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM     296    318   H4 (POTENTIAL).
FT   DOMAIN       319    370   CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM     371    390   H5 (POTENTIAL).
FT   DOMAIN       391    474   EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM     475    495   H6 (POTENTIAL).
FT   DOMAIN       496    686   CYTOPLASMIC (POTENTIAL).
FT   NP_BIND      483    605   CGMP (POTENTIAL).
FT   BINDING      542    542   CGMP (POTENTIAL).
FT   BINDING      557    557   CGMP (POTENTIAL).
FT   CARBOHYD     421    421   N-LINKED (GLCNAC).
FT   VARIANT       28     28   R -> Q.
FT                             /FTId=VAR_009295.
FT   VARIANT      114    114   D -> N.
FT                             /FTId=VAR_009296.
FT   VARIANT      316    316   S -> F (IN ARRP).
FT                             /FTId=VAR_009297.
FT   CONFLICT      46     46   S -> Y (IN REF. 1).
FT   CONFLICT      85     85   L -> I (IN REF. 1).
FT   CONFLICT     146    147   EE -> HH (IN REF. 1).
FT   CONFLICT     539    539   Y -> T (IN REF. 1).
FT   CONFLICT     677    678   GA -> WS (IN REF. 1).
SQ   SEQUENCE   686 AA;  79126 MW;  E520D0216C97AF6 CRC64;
```

```
Alignment Scores:
Pred. No.:
Score:                    1.47e-09
Percent Similarity:       412.00
Best Local Similarity:    40.60%
Query Match:              22.38%
DB:                       6.77%
                          1

Length:        686
Matches:       156
Conservative:  127
Mismatches:    288
Indels:        126
Gaps:          21
```

```
US-09-965-830-1_COPY_6_3257 (1-3252) x CNG1_HUMAN (1-686)

Qy   181 ATGCAGCGGGGGCTGTGCCCTGCCTTCTCCTTCCTTTATGGGCCAGACACCAGTGAGCTCCGGC 240
             |||||||                         |||||
Db    30 MetGluAsnGlyAlaCysSerSerPheSerGluAspSerAlaSerThrSerGlu 49

Qy   241 CAACAGGATCCGCAAGGCCCTGGACACGACACAGAGGACACACAAGGAGTTCAAGGGTCGAGCTGATCCTGTAC 300
             :::
Db    50 GluSerGluAsnGluAsnProHisAlaAlaArgGlySerPheSerThrArgSerLeu 67

Qy   301 CGGAAGAGCGGGCGTCCCGTTCTGGTTGTCTCCTGGATGTGATGATACCCATAAAGAATGAGAAA 360
         |||||||                                   |||   |||:::||
Db    68 ArgLysGlyGly                ProSerGlnArgGluGln 77

Qy   361 GGGGAGGTGGCTCTCTCCTAGTCTCTTCACAAGGACGATCACCGAAAACCAAG 411
             ::|||:||:::       |||    |||   |||
Db    78 TyrLeuProGlyAlaIleAlaIleAlaLeuPheAsnValAsn 89

Qy   412 AACCGAGGGGGCCCCGACAGTGAAGAGGAAGACACAGGTGGTGGGGCCCGGGGGCCCGGA 471
         |||            :::
Db    90 AsnSerSerAsnLysAspGlnLeuProGluGluLysLysLysGluGluVal 109

Qy   472 GCACGATCCAAAGGCTTCAAATGCCAACCAGGCGGGCGGGGGGCCGTGCTCTACCACCTG 531
         :::::||                       |||               ||||||
Db   110 SerLysSerAspAspLysAsnLysAsnAspPro 122

Qy   532 TCCGGGCCACCTGCAGAAGCGAGCCCCAAGGCACAGCTCAATAAGGGGGGTGTTTGGG 591
         |||                 |||   |||   |||
Db   123 GlyLysGlyLysLysLysLysAspAspLysGlyLysLys 134

Qy   592 GAGAAACCAAACTTGCCTGAGTACAAAGTAGCGCCCATCGGAAGTCGGCCCTTCATCCTG 651
             :::                          |||
Db   135 LysLysGluGluSerLysAspLysLysGluGluValValValValIle 154

Qy   652 TTGCACTGTTGGGCCACTGAGAGCCCACCTGGATGGCTTCATCCTGCTCGGCCACACTCTAT 711
                                        ||| ||| :::
Db   155 AspProSerGlyAsnThrTyrTyrAsnTrpLeuPheCysIleThrLeuProValMetTyr 174

Qy   712 GTGGGCTGTCACTGTCCCTACACGGCGTGTGTGAGCAGACACGAGGACCCCAGTGCCGCGC 771
         |||       :::
Db   175 AsnTrpThrMetValIleAlaArgAlaCysPheAspGluLeuGlnSerAspTyrLeuGlu 194

Qy   772 CGCGGCGCCGCCCAGCGGTGTGACCTGCCGTGGGAGTCCTCTCTTCATCCTTGACATTGTG 831
             :::                          :::
Db   195 ----LeuIleLeuAspTyrValSerAspIleValTyrLeuLeuAspMetPhe 212

Qy   832 CTGAATTTCCGTACCACATTCGTGTCCAAGTCGGGCGTGTTGCCCCAAAGTCC 891
             :::                          |||
Db   213 ValArgThrArgThrGlyTyrLeuGluGlnGlyLeuLeuValLysGluGluLeu 232

Qy   892 ATTTGCCTCACTACGTCACCACCTGTTCTGCTGGAGCTGAAGGGTGCCCTTT 951
         |||  ||||||||||:::
Db   233 IleAsnLysTyrLysSerAsnLeuGlnPheLysLeuAspValSerLeuLeuProThr 252

Qy   952 GACCTGCTACAATGCCTTCAAGGTCAC--GTGTACTTCGGGGCCATCGGGAGAGACG 1008
         ||||||||||:::::                       |||:::
Db   253 AspLeuLeuTyr---PheLysLeuGlyTrpAsnTyrProGluLeuArgLeuAsnArgLeu 271

Qy  1009 GTGCGCCGCTGCTTCGGCGCCCTGCTCGGGACCGGTACTCCGAGTACAGC 1068
         :::                       |||    |||
Db   272 LeuArgPheSerArgMetPheGluPhePheGlnArgThrGluThrArgThrAsnTyrPro 291
```

```
Qy  1069 GCCGTGGGTG---CTGACACTGCTTCATGCCGTTGTTCGCGCTGGGTCGCC 1125
              |||||||   |||
Db   292 AsnIlePheArgIleSerAsnLeuValMetTyrIleValIleIleIleHisTrpAsnAla 311

Qy  1126 TGCTTCTGGTTTTACATTGGCCGGAGAGATCGAGAGGAAGTCC---AGCTG---- 1179
              |||||||:::||    ||| :::          :::::::      ::
Db   312 CysValPheTyrSerIleSerGlyLysAsnAspThrTrpValTyrPro 331

Qy  1180 ----CCTGAGATTGGCGGTGGCTGCCCGCGCGACTGGGAGACTCCC 1179
              |||||||         |||           ||||||||:::
Db   344                ArgLeuAlaArgLys 344

Qy  1228 TACTACCTGGTGGGGCGGAGGCCAGCTGGGAGGAACAGCTCCGGCGAGTGACAACTGC 1287
              :::::::|||||||:::               :::       ::: |||
Db   332 AspIleAsnAspProGluPheGly-------ArgLeuAlaArgLys 344

Qy  1288 AGCAGCAGCGAGGAGGCCAACGAGGGGACGGGGCTGGAGCTGGGAGGACTCC 1347
              :::   |||              |||                 :::
Db   344                          344

Qy  1348 AGCGGCCTACATCAACTCCCTACTTCGCCACTCAGCAGCCGTGGGGCTTCGGC 1407
              ||||||::::::   |||||||:::::::|||:::   :::::: |||
Db   344                          344

Qy  1408 AACGTGTCCGCCAACACGGACACCGAGAAGATCTTCTCCATCATGCACCATGCTCATCGGC 1467
              :::   |||:::|||               :::|||||||:::|||
Db   345 TyrValTyrSerLeuTyrTrpSerThrLeuThrLeuThrThrIleGly---Glu 361

Qy  1468 GCCCTGAATGCACGCGGCGGTGGTGTTGGGAACGTGACAGCCCATCATCCAGCGCGATGCTACGGCC 1527
              :::   |||   |||:::||||||:::               |||:::
Db   362 ThrProProProValArgAspSerGluTyrValPheValValAlaAspPheLeuLeuIleGly 381

Qy  1528 CGGCCGCTTTCTGTACCACAGCCGCACCGCGGGGCAGCTGGCCCGGCCCGC 1587
              :::|||            |||:::|||               :::|||
Db   382 ValLeuIlePheAlaThrIleValGlyAsnIleGlySerMetIleSerAsnMetAsnAla 401

Qy  1588 ATCCCACAAGCCCCTCAAGCAGCCCATGCTGGAGTACTTCCAGGCACCACTGGGCGGTGAAC 1647
              :::|||               |||:::|||||||:::|||||| :::
Db   402 AlaArgAlaGluPheGlnAlaArgIleAspAlaIleLysGlnTyrMetHisPheArgAsn 421

Qy  1648 AATGGACATCGACACCACCGACGTGCTCAGCAGACGCCTCCCTGACGAGCTGCCGCCGAGACATC 1707
              :::   |||:::|||               |||:::|||:::|||||| :::
Db   422 ValSerArgAspMetGluLysArgValIleLysTrpPheAspTyrLeuTrpThrAsnLys 441

Qy  1708 GCCCATCCACCTCACACAAGGACGATGCCCTGCAGGCCCTGACTTTTGTGTCTCGGCTCCATGGAG 1764
              :::|||            :::::: ||| ||| :::|||:::|||
Db   442 LysThrValAspGluLysGluValLeuLysTyrLeuProAspLysLeuArgAlaGluIle 461

Qy  1765 GGCTGCCTGCGGGCCGACTGTCTCCTGGGCCCGCCTTCTGCGGCCCGGCGAGTAC 1824
              |||:::||||||||| :::           |||:::|||||||
Db   462 AlaIleAsnValHisLeuAspThrLeuLysLysValArgIlePheAlaAspCysGluAla 481

Qy  1825 CTCCATCCACCAAGGGACGATGCCCTGCAGGCCCTGACTTTTGTGTCTCGGCTCCATGGAG 1884
              :::|||            |||::: ||| :::||||||:::|||
Db   482 GlyLeuLeuValGluLeuValLeuLysLeuGlnProGlnValTyrSerProGlyAspTyr 501

Qy  1885 GTGGCTC-----AAGGGTGGCACCGTGCTGCTCCATCCTTAGGGAAGGGCGACCTGATCGGC 1938
              |||       |||::::::|||               |||:::
Db   502 IleCysLysLysGlyAspIleGlyArgGluMetTyrIleIleLysGluGlyLysLeuAla 521

Qy  1939 --------TGTGAGCTGCCCCGGGGCAGGTGGTGGAAAGGCCAATGCCGACGTG 1986
              :::||| |||:::||||||               |||:::
Db   522 ValValAlaAspAspGlyValThrGlnPheValValLeuSerAspGlySerTyrPheGly 541

Qy  1987 AAGGGGCTGACTGACGTACTGCCGTCCTCCAGCGTGCTGCCTGACGACGCCTT 2046
              |||         |||:::|||               |||::::::|||:::
Db   542 GluIleSerIleLeuAsnIleLysGlySerLysAlaGlyAsnArgArgThrAlaAsnIle 561

Qy  2047 GCGCTGGTTGCCCCAGAGTTTGCCCCGCGCTTC----AGTCGTTGGGCCTCCGGA 2091
              |||||||:::               :::|||   ||| :::
Db   562 LysSerIleGlyTyrSerAspLeuPheCysLeuSerLysAspAspLeuMetGluAlaLeu 581

Qy  2092 ---GGGGAGCGCTCAGCTACACCACCTGGGTTCCAACCTGGGTGGGGGAGGCTCTGCAGAGGTTGGCTCTGCGCGAGGAGGTTGGACACCAGC 2148
              |||   |||               |||:::               |||:::
Db   582 ThrGluTyrProAspAlaLysThrMetLeuGluGluLysGlyLysGlnIleLeuMetLys 601

Qy  2149 TCCCTGAGCGGCGGACAATACCCTTATGTCCACGCTGGAGGAGAGAGACA 2199
              |||||| |||               |||:::
Db   602 AspGlyLeuLeuAspLeuAsnIle----AlaAsnAlaGlySerAspProLysAsp------ 618

Qy   619 -------------------- 619
              |||:::||||||         |||
Db   619 LeuGluGluLysValThr 624
```

RESULT 14
CNG3_CHICK

```
ID   CNG3_CHICK     STANDARD;      PRT;    645 AA.
AC   Q90980;
DT   01-NOV-1997 (Rel. 35, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   01-NOV-1997 (Rel. 35, Last annotation update)
DE   Cyclic nucleotide gated channel, ROD photoreceptor, alpha subunit
DE   (CNG channel 3) (CNG-3) (CNG3).
OS   Gallus gallus (Chicken).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC   Gallus.
OX   NCBI_TaxID=9031;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=93264082; PubMed=7684234;
RA   Boenigk W., Altenhofen W., Mueller F., Dose A., Illing M.,
RA   Molday R.S., Kaupp U.B.;
RT   "Rod and cone photoreceptor cells express distinct genes for
RT   cGMP-gated channels.";
RL   Neuron 10:865-877(1993).
CC   -!- FUNCTION: VISUAL SIGNAL TRANSDUCTION IS MEDIATED BY A G-PROTEIN
CC       COUPLED CASCADE USING CGMP AS SECOND MESSENGER. THIS PROTEIN CAN
CC       BE ACTIVATED BY CYCLIC GMP WHICH LEADS TO A OPENING OF THE CATION
CC       CHANNEL AND THEREBY CAUSING A DEPOLARIZATION OF ROD
CC       PHOTORECEPTORS.
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC   -!- SIMILARITY: BELONGS TO THE CYCLIC NUCLEOTIDE-GATED CATION CHANNEL
CC       FAMILY.
CC
CC   --------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------
DR   EMBL; X89599; CAA61758.1; -.
DR   InterPro; IPR000636; M+channel_nlg.
DR   InterPro; IPR000595; cNMP_binding.
DR   Pfam; PF00027; cNMP_binding; 1.
DR   Pfam; PF00520; Ion_trans; 1.
DR   SMART; SM00100; cNMP; 1.
DR   PROSITE; PS00888; CNMP_BINDING_1; 1.
DR   PROSITE; PS00889; CNMP_BINDING_2; 1.
DR   PROSITE; PS50042; CNMP_BINDING_3; 1.
KW   Ionic channel; Ion transport; cAMP-binding; Transmembrane; Vision;
KW   Multigene family;
FT   DOMAIN        1    117       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    118    137       H1 (POTENTIAL).
FT   DOMAIN      138    150       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    151    169       H2 (POTENTIAL).
FT   DOMAIN      170    193       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    194    213       H3 (POTENTIAL).
FT   DOMAIN      214    251       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    252    274       H4 (POTENTIAL).
FT   DOMAIN      275    326       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    327    346       H5 (POTENTIAL).
FT   DOMAIN      347    430       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    431    451       H6 (POTENTIAL).
FT   DOMAIN      452    561       CYTOPLASMIC (POTENTIAL).
FT   NP_BIND     439    561       CAMP (BY SIMILARITY).
FT   BINDING     498    498       CAMP (POTENTIAL).
```

```
FT   BINDING        513    513     CAMP (POTENTIAL).
FT   CARBOHYD       377    377     N-LINKED (GLCNAC...) (POTENTIAL).
SQ   SEQUENCE       645 AA;  74778 MW;  989D51F61AC7D31 CRC64;

Alignment Scores:
Pred. No.:
Score:                   1.62e-09
                         411.00
Percent Similarity:       43.83%
Best Local Similarity:    23.61%
Query Match:               6.75%

US-09-965-830-1_COPY_6_3257 (1-3252) x CNG3_CHICK (1-645)

                         Length:        645
                         Matches:       132
                         Conservative:  113
                         Mismatches:    240
                         Indels:         74
                         Gaps:           17
                         DB:              1

Qy   547  AAGCAGCCCAAGGGCAAGCAAAGCTCAATAAGGGGTGTTGGGGAGAAACCAAACTTG 606
          |||...||   |||||||||||||||||||  |||||||||  |||||||||||||
Db    84  LysGluLysLysGluLysHisLys---AsnLys------AsnLysLysGlyLys 99

Qy   607  CCTGAGTACAAAGTAGGCGCCATCCGGAAGTCGGCCCTTCATCCTGTTGCACTGTGGGGCA 666
          ||:||||||   |||||   ||||||||||||||||||||||  ||||||||||||
Db   100  GluGluGluLys------------LysLysAspIlePheIleIleAspProAlaGlyAsn 115

Qy   667  CTGAGAGCCACCTGGGATGGCTTCATCCTGCTCGCCACACTCTATGTGGCTGTCACTGTG 726
          :::||||||   |||||||||   |||   ||||||||||||||   :::||||||||
Db   116  MetTyrTyrAsnTrpLeuPheCysIleThrMetProValMetTyrAsnTrpPheThrMetIle 135

Qy   727  CCCTACACGGTGTGTGGAGCACAGCAGGGAGCCCAGTGCCGCCCGCCCAGC 786
          |||   |||   |||   |||   |||   |||||||||||||||
Db   136  IleAlaArgAlaCysPheAspGluLeuGlnAsnAspTyrLeuAlaValTrp------Phe 153

Qy   787  GTCTGTGACCTGGCCGTCGGAGGTCCTTCTACATTGTCTTGACAATTGCTGAATTTCCGTACC 846
          ||||||||   ||||   :::|||   :::|||   |||   ||||||
Db   154  IleValAspTyrValSerAspValIleTyrIleAlaAspMetPheValArgThrArgThr 173

Qy   847  ACATTCGTACTGTCCAAGTCGCGGCCAGGTGGTGTTGCCCAAAGTCCATTTGCCTCACTAC 906
          |||   |||   |||||||||||   |||||||||||   :::|||||
Db   174  GlyTyrLeu---GluGlnValLeuLeuValLysGluGlnLysGluLysTyr 192

Qy   907  GTCACCACC---TGGTTCCTGCTGGAGTGTCATCCGGACGCGTGCCCCTTGACCTGCTACAT 963
          |||||   :::   |||||||   ||||||||   |||||   :::|||
Db   193  LysSerSerLeuGlnPheLysLeuAspPheLeuSerIleLleProIleThrAspLeuLeuTyr 212

Qy   964  GCCTTCAACGTCAACGTG---TACTTCGGGGGGCCCATWCTGCTGAAGACGGTGCGGCCCTGCTG 1020
          :::|||   |||||||||   :::|||||||   |||||||||||:::
Db   213  ---PheLysLeuGlyLeuAsnTyrProGluLeuAsnArgLeuLeuArgValAla 231

Qy   1021 CGGCCTGCTGCGGCCCTGCCTTCCGCGCGGTGACCGGTACTCCAGTACAGCGCCTGGTG--- 1077
          |||   |||   :::|||   |||||||   |||||||:::
Db   232  ArgMetPheGluPhePheGlnArgThrGluThrArgThrAsnTyrProAsnIlePheArg 251

Qy   1078 CTGACACTGCTCATGGCCGTGTTCGCCCTGCTGCCCACTGGGGTGCCCTGCGTCTGGTTT 1137
          ||||||||||||||||||||||   ||||||||||||||
Db   252  IleSerAsnLeuValMetTyrIleValIleIleIleIleHisTrpAsnAlaCysVal--- 269

Qy   1138 TACATTGGCCAGGGGGAGATCGAGAGCACGCAACCGACGCTGCCTGAGAATTGGCTGGCTG 1197
          ----------------------------------------------------
Db   269  ------------------------------------------- 269

Qy   1198 CAGGAGCTGGCCCCGCCGCACCTGGAGACTCCCTACCTGGTCGGCCGGAGGCCAGCTGGA 1257
          |||||||
Db   270  ------------------------------TyrTyrSerIleSerLysAlaIleGlyPhe 279

Qy   1258 GGGAACAGCTCCGGCCCAGAGTGACAACTGCAGCAGCAGCGAGCCAAACGGACGGGG 1317
          |||   |||
Db   280  GlyAlaAspThrTrpAsnThrSerHisProGluPheAlaArg------------ 296

Qy   1318 CTGGAGCTGCTGGGCGGCCCCGCCGTGCTGCGCCAGCCCTTCACTACCTCCTCTACTTCGGCA 1377
          |||
Db   297  ------------LeuThrArgLysTyrValTyrSerLeuTyrTyrSer 308

Qy   1378 CTCAGCAGCCTCACCAGCGTGGGCCTTCGGCAACAGTGTCCGCCAACACGACACCGAGAAG 1437
          :::||||||||||:::||:||
Db   309  ThrLeuThrLeuThrThrIleGly------GluThrProProProValArgAspSerGluTyr 327
```

```
RA   Flockerzi V., Dascal N., Kaupp U.B., Hofmann F.;
RT   "Primary structure and functional expression of a cyclic nucleotide-
RT   gated channel from rabbit aorta.";
RL   FEBS Lett. 329:134-138(1993).
CC   -!- FUNCTION: ODORANT SIGNAL TRANSDUCTION IS PROBABLY MEDIATED
CC       BY A G-PROTEIN COUPLED CASCADE USING CAMP AS SECOND MESSENGER.
CC       THE OLFACTORY CHANNEL CAN BE SHOWN TO BE ACTIVATED BY CYCLIC
CC       NUCLEOTIDES WHICH LEADS TO A DEPOLARIZATION OF OLFACTORY
CC       SENSORY NEURONS.
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC   -!- SIMILARITY: BELONGS TO THE CYCLIC NUCLEOTIDE-GATED CATION CHANNEL
CC       FAMILY.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; X59668; CAA42201.1; ALT_INIT.
DR   InterPro; IPR000636; M+channel_nlg.
DR   InterPro; IPR000595; CNMP_binding.
DR   Pfam; PF00027; cNMP_binding; 1.
DR   Pfam; PF00520; Ion_trans; 1.
DR   SMART; SM00100; cNMP; 1.
DR   PROSITE; PS00888; CNMP_BINDING_1; 1.
DR   PROSITE; PS00889; CNMP_BINDING_2; 1.
DR   PROSITE; PS50042; CNMP_BINDING_3; 1.
KW   Ionic channel; Ion transport; cAMP-binding; Transmembrane;
KW   Multigene family; Olfaction.
FT   DOMAIN        1    140       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    141    160       H1 (POTENTIAL).
FT   DOMAIN      161    173       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    174    192       H2 (POTENTIAL).
FT   DOMAIN      193    216       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    217    236       H3 (POTENTIAL).
FT   DOMAIN      237    274       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    275    297       H4 (POTENTIAL).
FT   DOMAIN      298    349       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    350    369       H5 (POTENTIAL).
FT   DOMAIN      370    453       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    454    474       H6 (POTENTIAL).
FT   DOMAIN      475    664       CYTOPLASMIC (POTENTIAL).
FT   NP_BIND     462    584       CAMP (BY SIMILARITY).
FT   BINDING     521    521       CAMP (POTENTIAL).
FT   BINDING     536    536       CAMP (POTENTIAL).
FT   CARBOHYD    379    379       N-LINKED (GLCNAC...) (POTENTIAL).
SQ   SEQUENCE   664 AA;  76205 MW;  5E917D0B322B3E9 CRC64;
```

US-09-965-830-1_copy_6_3257 (1-3252) x CNG2_RABIT (1-664)

Alignment Scores:
Pred. No.:
Score: 405.00
Percent Similarity: 38.84%
Best Local Similarity: 23.95%
Query Match: 6.65%
DB:

2.75e-09

Length:	664
Matches:	177
Conservative:	110
Mismatches:	262
Indels:	190
Gaps:	29

```
Qy   343 CCCATAAAGAAATGAGAAGGGAGGTG---GCTCCTCTTCCCTAGTCCTTCCACAAGGACATC 399
             |||                   :: :               |||||
Db    12 ProAlaAsnAsnHisAsnHisValProAlaThrIleLysAlaAlaAsnGlyLysAspGlu 31

Qy   400 AGGGAAACCAAGAACCGA------ 417
             |||||:::||::::|||
Db    32 SerArgThrArgSerArgProGlnSerAlaAlaAspSerAspThrSerGluLeuGln 51

Qy   418 ------GGGGGCCCGACAGATGGAAGGAGAGACAGGTGGTGGCCGGCGCCGATATGGCCGG 471
                   ::::: |||||||
Db    52 ArgLeuLeuAlaGluMetAspAlaProGlnGlnArgArgGlyGlyPheArgArgIleValArg 71

Qy   472 GCA------...CGATCCAAAGGCTTCAATGCCAACCGCGGCGGGGAGC 510
             |||                              |||
Db    72 LeuValGlyValIleArgGlnTrpAlaLeuArgAsnPheArgGluGlnValLeuArgPro 91

Qy   511 CGGGCCGTGCTGCTCTACCACCTGTCCGGG---CACCTGCAG------AAGCAGCCC 555
             :::         |||||||          :::||||
Db    92 AspSerPheLeuGluArgPheArgGlyProLeuGlnThrValThrThrGlnGlyGly 111

Qy   556 AAGGGCAAGCAAGCTCAATAAGGAGGGTGTTTTGGGAGAAACCAAAACTTGCCTGAGTAC 615
             |||||||                |||
Db   112 AspGlyGlyLys------...------GlyAspGlyAspGly------AspGly 121

Qy   616 AAAGTAGGCCGCCATCCGGAAGTCGCCCTTCAICCTGTTCACGTGGGGCACTGAGGCC 675
             :::    |||              |||
Db   122 LysGlyGlyThrLysSerLysPheGluValLeuAspAspProAlaGlyAspPro... 141

Qy   676 ACCTGGGATGGCTTCATCCTGCTCGCCACACTCTATGGGGTGTCACTGCCCTACAGC 735
             |||                ::: |||||            ::: |||
Db   142 ArgTrp------...------LeuPheValIleAlaMetProLeuTyrAsn 154

Qy   736 GTGTGTGTGGAGCAGCAGCACGGGAG------CCCAGTGCCGCCGGGGC------CCG 780
             |||||||                       ||||||
Db   155 TrpCysLeuLeuValAlaAlaArgAlaCysPheSerAspLeuGlnArgGlyTyrPheLeuVal 174

Qy   781 CCCAGCGTCTGTGACCTGGCCGGTGGAGGTCCTCTTCATCCTTGACATTGTCGTGAATTTC 840
             :::      |||||            ::::||||             ||| |||::
Db   175 TrpLeuValLeuAspProTyrPheSerAspValValTyrIleAlaAspLeuPheIleArgLeu 194

Qy   841 CGTACCACATTCGTGCCAAGTCGGGCGCCAGGTGGTGTTTCCCCAAAGTCCATTTGCCTC 900
             ::::||||                           ::: |||
Db   195 ArgThrGlyPheLeu---GluLeuLeuValLysAspProLysSerLeuArgAsp 213

Qy   901 CACTACGTC----ACCACCTTGGTTCCCTGCGATGTCAGTCAGCGGCGCCCTTTGACCTG 957
             ::::||||||                    ||| |||:::
Db   214 AsnThrTyrIleHisThrLeuGlnProPheLysLeuLeuAspGlnSerIleIleLeuPro 233

Qy   958 CTACATGCCTTCAAGGTCAACGTGTACTTC-----GGGCCCAAT-----CTGCTGAAG 1005
             |||                ||||||       |||:::
Db   234 IleTyrPheAlaValGlyIleHisSerAsnProGluLeuArg 246

Qy  1006 ACGGTGCCGCCCTGCTCGCGCCCTGCCTTCCCGGGCTTTGCTGGAGCCGGGTACCTCG---- 1059
             |||||||                         ::: |||
Db   247 PheAsnArgLeuLeuHisPheAlaArgMetPheGluPhePheAspArgThrGluThrArg 266

Qy  1060 ------CACTACACGGCCGTGGTCGTGCTGACAGTCGTCATGGCCGTG 1098
                   ||| :::                |||||||:::
Db   267 ThrSerTyrProAsnIlePheArgIleSerAsnLeuValLeuTyrIleLeuLeuIleIle 286

Qy  1099 TTCGCCCTGCTCGCCCACTGGGTCGCTCGTGGTTTTACATTGGCCACGACGGGAGATC 1158
             ||||:::          ||| :::||||              ||| :::
Db   287 HisTrpAsnAlaCysIleLeuTyrTyrTyrAlaIle 296

Qy  1159 GAGAGCAGCGAAATCCCGAGCTGCCTGAGAATTGGCTGGCTGCAGGAGCGCCGGCGCCGACTG 1218
             ::: |||              :::||||
Db   296 ------...------ 296

Qy  1219 GAGACTCCCTACTACCTGGTGGGCCCGGAGGCCAGCTGGAGGGGAACAGAGCTCCGGCCAGAGT 1278
             ||||::                  |||
Db   296 ------...------ 296

Qy  1279 GACAACTGCAGCAGCAGCAGGCCAAGGGACGAGGGGGGTGGAGCTGCTGGGGCGGCCCG 1338
             |||:::::                   |||
Db   297 SerLysSerIleGlyTyrGlyValAspThrTrpValTyrPro 310

Qy  1339 TCG------CTCGGCAGCGCGGCCTACATCACCTCCTCTACTTC 1374
             :::              |||||||:::
Db   311 AsnIleThrAspProGluTyrGlyTyrLeuAlaArgGluTyrTyrCysLeuTyrTrp 330

Qy  1375 GCACTCAGCAGCCTCACCAGCTGGGGCTTCGGCAACGTGCCCAACACGGGACACCGAG 1434
             |||:::::              ||| :::
Db   331 SerThrLeuThrLeuThrThrIleGly---GluThrProProValLysAspGluGlu 349

Qy  1435 AAGACTCTTCCCATCTGCACCATGCTCACGGCGCCCTGATGCACAGCGGGTTGGGTGTTTGGG 1494
             :::::||||              |||
Db   349 ArgLeuAlaLeuAlaGluMetAspAlaProGlnGlnArgArgGlyGlyPheArgArgIleValArg 71
```

```
Db   350 TyrLeuPheValIlePheAspPheLeuIleGlyValLeuIlePheAlaThrIleValGly 369
            :::||||||     :::||||||       :::|||     :::|||        :::|||
Qy  1495 AACGTGACGGCCATCACCGCAGTGTACGCCCGCGCCTTCTGTACCACAGCCCGGACG 1554

Db   370 AsnValGlySerMetIleSerAsnMetAlaThrArgAlaGlyPheGlnAlaLysIle 389
            ||||||||      :::|||       ||     :::|||
Qy  1555 CGCGACCTGGGCGACTACATCACCGCCATCCCAAGCCCCTCAAGCCAGCGCATG 1614

Db   390 AspAlaValLysHisTyrMetGlnPheArgLysValSerLysGlnMetGluAlaLysVal 409
            ||||||||||      :::|||::::::
Qy  1615 CTGGAGTACTTCCAGGCCACCACCTGGCGGTGAACAATGGCATCGACACCACCGAGCTGCTG 1674

Db   410 IleLysTrpPheAspTyrLeuTrpPheThrAsnLysLysThrValAspGluArgGluValLeu 429
            ||            :::|||           :::||||||      ||||||||
Qy  1675 CAGAGCCTCCCTGACGAGCGTGCGCGCAGACATGCCACCTGCCACAAGGAGGTCCTG 1734

Db   430 LysAsnLeuProAlaLysLeuArgAlaGlyIleLeuAlaIleLeuAsnValHisLeuSerThrLeu 449
            :::|||||||:::|||        ||      :::||:::::::::|||
Qy  1735 --CAGCTGCCACTGTTTGAGGCGGCCACCGCCCTGCCGGGCCACTGTCTCTGGCC 1791

Db   450 LysLysValArgIlePheGlnAspCysGluLeuAlaGlyLeuLeuGluLeuValLeuLys 469
            :::|||      :::|||::      ||           ||||||
Qy  1792 CTGCGGCCCGGCCTTCTGCACGCCGGGAGTAACCTCATCACCAAGGCCGATGCCCTGCAG 1851

Db   470 LeuArgProGlnValPheSerProGlyAspTyrIleCysArgLysGlyAspIleGlyLys 489
            :::||||||||      :::|||
Qy  1852 GCCCTCTACTTTGTCTGTCTGTCTCGGCTCCATGGAGGTGCTC---- 1905

Db   490 GluMetTyrIleIleLysGlyLysLeuAlaValValAlaAlaSpAspGlyValThrGln 509
            |||        :::|||        |||
Qy  1906 CTCGGCCATCCTAGGGAGAGGGGCGACCTGATCGGC-------- 1953

Db   510 TyrAlaLeuLeuSerAlaGlySerCysPheGlyGluLeuSerIleLeuAsnIleLeuLysGly 529
            :::                     :::
Qy  1954 CGGGAGGAGGTGGTAAAGGCCAATGCCGACGTGAAAGGGGCTGACGTACTGCGTCCTGCAG 2013

Db   530 SerLysMetGlyAsnArgArgThrAlaAsnIleLeuArgSerLeuGlyTyrSerAspLeuPhe 549
            ||||||     :::||||||::::::|||
Qy  2014 TGTCTCCAGCGTGGCTGGCCTGCCGACGACAGCCCTTGCGCCTGTACCCCGAGTTTGCCCCGC 2073

Db   550 CysLeuSerLysAspAspAspMetGluValAlaValThrGluTyrProAlaLysLysVal 569
            ||            :::::::::
Qy  2074 TTC---AGTCGTGGC----------CTCCGAGGGGACTCACCTACAACCTGGGT 2115

Db   570 LeuGluGluArgGlyArgGluIleLeuLeuMetLysGluLeuLeuAspGluAsnGluVal 589
            :::||||||        ||         |||
Qy  2116 GCTGGGGGAGGCTCTCGCAGAGGTGGACAC---AGCTCCCTGAGCGGCGACAATACCCTT 2172

Db   590 AlaSerMetGluValAspValGlnGluGlyLysLeuLysGluThrAsn 607
            |||||||||||
Qy  2173 ATGTCCACGGTG---------- 2184

Db   608 MetGluThrLeuTyrThrArgPheGlyArgLeuLeuAlaGluTyrThrArgGlyValAlaGlnGln 627
            |||||||
Qy  2185 ---------------GAGGAG 2190

Db   628 LysLeuLysGlyIleThrValLeuGluGlyValLysMetLysGlnAsnThrGluAspAsp 647
                                      :::|||
Qy  2191 AAGGAGACAGAGTGGGGAGCAGGGCCCCACGGTCTCCCAGCCACCTGATGAGCCC 2247

Db   648 TyrLeuSerAspGlyMetAsnSerPro------GluProAlaAlaAlaGluGluGlnPro 664
            :::|||||||      |||::::::|||
```